A GIS-SUPPORTED ENVIRONMENTAL RISK ASSESSMENT FOR PETROLEUM WASTE CONTAMINATED SITE

Su Chen1, Gordon Huang1, and Jonathan Li2

1Environmental System Engineering Program, Faculty of Engineering, University of Regina, Regina, SK, Canada S4S 0A2
Tel: (306) 585-5531, Fax: (306) 585-4095, E-mail: chen11su.gordon.huang@uregina.ca
2Department of Geography, University of Regina, Regina, SK, Canada S4S 0A2
Tel: (306) 585-5273, Fax: (306) 585-4815, E-mail: jun.li@uregina.ca

KEYWORD: GIS, environmental risk assessment, Monte Carlo, fuzzy set theory.

ABSTRACT

Responding to an urgent requirement for effective management and environmental assessment of contaminated petroleum sites, this paper presents an environmental risk assessment approach for petroleum-contaminated aquifers due to leakage from underground storage tanks (USTs). It contains two components: environmental risk assessment and geographic information system (GIS). Using the monitoring data and predicted fate of contaminants concentration under different transportation time and different proportion of remediation strategies groundwater transportation model software-kit, the risk assessment model can be effectively incorporated with effects of different contaminants and different remediation techniques within one framework. Extensive uncertainties for a number of modeling inputs in an exposure-dose model are presented as a series of probability distribution function with a variety of forms, and are solved using the classic Monte Carlo simulation method. Data originating from the exposure dose model are sequentially incorporated into a fuzzy risk assessment approach to achieve an integrated risk level for the target location. Results of a case study indicated that plausible solutions for risk assessment under different system conditions have been generated.

GIS provides comprehensive database of contaminated site conditions, tools for spatial and customized interface of risk assessment, and visual presentation of modeling results and site natural and spatial characterization. Especially, integration of the risk assessment results with Geospatial information is very helpful for identifying and assessing pollution impacts on specific receptors. A desktop ArcView GIS was used in this study for spatial analysis and Microsoft Excel and Access were used for the spreadsheet and relational database components, respectively.

1. INTRODUCTION

Development of the petroleum industry is currently associated with a number of environmental concerns. Among them, severe soil and groundwater contamination phenomenon is attracting more and more attentions from the public, governments, and petroleum industries themselves (Mathews and Donahue 1989). This situation is especially worse in western Canada where involves extremely active petroleum production, processing, upgrading procedure. Showing by the past and recent research, it is recognized that the major sources causing soil and groundwater contamination are coming from leaking storage tanks which are used enormously by commercial, industrial and residential sectors nationwide.

In North America, several hundred thousand of USTs that are used for storing petroleum products are leaking (Levy et al., 1990). There are over 70 separate hydrocarbon compounds in regular gasoline (Bruesl and Hoag, 1984), and they are mainly compounds of aliphatic hydrocarbons and aromatic hydrocarbons.

The greatest harm caused by a leaking underground storage tank, which holds petroleum or petroleum by-products, is the contamination of groundwater. Petrochemical compounds that seep down to a groundwater formation will intent to float on top of the water table according to their lighter specific gravity. Also, volatile components can exist in gaseous phase and will escape as fumes or odors. Other components, such as benzene, toluene, and various xylenes (or BTXs), can attach to the soil and exist in adsorbed phase. In addition, some compounds can be in soluble phase contained within the groundwater (Canter et al., 1999).

Soil and groundwater contamination can lead to a variety of impacts, risks, and liabilities to the communities and for the industries themselves. The leakage represents an increasing danger to groundwater resources and public health (Testa and Winegardner, 1991; Hayward, 1994). For example, it is proved that one gallon of gasoline can render one million gallons of water unsuitable for drinking needs. Since the nation draws about half of its drinking water from groundwater sources, there is growing concern that leaking underground storage tanks will continue to contaminate many drinking water sources. Exposure to contaminated soil and groundwater can occur through skin contact, inhalation, or ingestion. Very truly, even a small leak into an underground water table can be permanently damaging to the source, since groundwater is unable to naturally recharge and cleanse itself because petroleum and its bi-products float on top of the water. The leakage problem has led to a variety of impacts, risks, and liabilities. In Canada, about 10% of 200,000 underground storage tanks are leaking and contaminating the surrounding environment, causing violent losses of thousands of dollars annually to petroleum industries and stakeholders. It is estimated that rendering all of these abandoned drilling stumps will require a minimum expenditure of 10 billion dollars, what a shock? Therefore, in-depth and effective environmental risk assessment of groundwater contamination due to leaking petroleum contaminants is important and anxiously desired for evaluating the need for site remediation actions and providing support for decisions related to prevention, estimation, and remediation of the leakage and contamination problems (Huang et al., 1999).

Since 1970 the field of risk assessment has received widespread attention within both the scientific and regulatory communities and the legal system (Pausenbach, 1990). In recent years, risk assessment techniques have become widely utilized in the decision making process related to contaminated soil and groundwater problem, it could support managers with a more rational and scientific base on which kind of decision should be made.

Generally, the formulation of the environmental risk problem captures the entire process of identifying the source term of risk agent, toxicity assessment and exposure assessment and risk characterization. This process involves a number of chemical, physical, biological factors regarding to their direct or indirect relations to the environmental risk problem. The related parameters generally show high degrees of intrinsic variability and substantial levels of uncertainty since many system components in real-world problem many not be known with certainty (Woodbury et al., 1991). This makes the study...
systems more complicated 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). There have been a lot of studies to date to deal with the uncertainties within the environmental risk assessment for Petroleum Hazardous Waste concern. Most of the studies are employing probability theory (e.g., Monte Carlo simulation). For instance, Paustenbach (1999) proposed a comprehensive methodology for assessing the risks to human and wildlife posed by contaminated soil which involving Dioxin (1989). Adams and Hamra (1994) processed a health risk assessment using a Latin Hypercube probabilistic risk assessment of health risks associated with exposures to contaminated sediment and biota in an estuary.

Another major approach for uncertainty involved environmental risk assessment is through fuzzy set theory, which is suitable for situations where information is not available (uncertainties present as fuzzy membership functions rather than probability distribution functions) (Gardossy et al., 1991). For example, Lee and Dahab (1994) et al. developed a rule-based fuzzy set approach for risk assessment of nitrate-contaminated groundwater. Kangai and Riggs (1989) used fuzzy number and natural language to represent the uncertainty in construction risk assessment process. Huang (1995) proposed a fuzzy risk assessment method to assess the natural hazards in the urban area. Huang and Chen et al. (1999) developed a fuzzy relation approach to analysis the risk causing by the leachate from the UST.

However, results from evaluations of human health risks associated with environmental contamination are traditionally presented non-spatially. Non-spatial tabular reporting of single value has been well accepted convention for characterization of human health risk, no specific spatial information is included (William and Daniel, 1999). However, it is also very clear that risk assessments have an important spatial component, for the reason that risk evaluations are generated by reference to specific environmental health data collected from specific locations. Regarding to the above mentioned problems, GIS components can provide a comprehensive database of contaminated site conditions, tools for spatial and customized interface of risk assessment, and visual presentation of modeling results and site natural and spatial characterizations. Especially, integration of the risk assessment results with spatial land-use information will be very helpful for identifying and assessing pollution impacts on specific receptors through various exposure pathways, maps can be valuable for risk communication. At present, however, such integrated GIS and risk assessment applications are relatively rare and an important reason for this situation appears to be limited awareness among risk analysts of the full capacities of GIS technology.

As an extension to the previous research, a probabilistic and possibilistic approach is utilized to deal with environmental risk assessment concerning the petroleum-contaminated site. The process of risk assessment in this research involves the following steps:

- A 3D multiphase transport model (BIOF&T), a finite element model from Resources & Systems International, Inc., is preliminarily employed for studying the transport process of hydrocarbon-derived contaminants in soil and groundwater.
- Individual health risk characterization through widely accepted exposure-dose model, in this part Monte Carlo simulation method and fuzzy relation analysis are employed to evaluate the health risk posed by different age group.
- Site overall health risk assessment using a fuzzy interval risk assessment approach.
- Results visualization integrated with GIS.

2.1 Site Characteristics and Contamination Status

The research target, the Hoosier Site, is located at approximately 50 km northwest of Kindsley, Saskatchewan, Canada. Up to date, the facilities on site include an adsorption tower building, an above ground naphtha storage tank, a heater building, a cooler unit, a truck loading dock, a salt bath line heater, a flare pit, and above ground pipelines. They had been in operation as a natural gas processing plant ever since mid1960s to early1990s.

The Hoosier Gas Adsorption Plant (HGAP) was designed and utilized to clean natural gas through a series of scrubbers, removing the contaminants from gas by washing impurities. Throughout the history of the site operation, naphtha condensate, a waste liquid removed from the gas by a series of scrubbers, has been disposed of in an underground storage tank (UST) located immediately south of the adsorption tower building (Roper Environmental Engineering Inc., 1992; Clifton Associates Ltd., 1996).

Several studies have been conducted to evaluate the nature and extent of any permanent residuals at Hoosier Site. It is understood that the UST was perforated to accommodate naphtha disposal via ground infiltration. More recently, the naphtha condensate was disposed of in an above ground storage tank (AST) located approximately 32 m south of the adsorption tower building. Naphtha was trucked off the site and used to blend with heavy oil. The natural gas condensate was believed to spill and leak into soil following the seepage into the groundwater (Clifton Associates Ltd., 1994). Two contaminant-concentrated zones were then formed in the subsurface capillary zone (at the interface between unsaturated and saturated zones). Numberous of Environmental investigations of the Hoosier Site were conducted by Environmental Consulting Companies. Among them, including Roper Environmental Engineering Inc. in 1992, Clifton Associates Ltd. in 1993, 1995 and 1998, Stanley Consulting Group Ltd. in 1997, and Clifton Associates Ltd in 1998. These investigation works included surface and subsurface sampling of soil and groundwater, bore hole drilling, and monitoring well construction.

Results of the investigations identified the existence of free liquid phase contaminants free phase products in several monitoring wells (Clifton Associates Ltd., 1994, 1995, 1996; Stanley Consulting Group Ltd., 1997). Residual phase hydrocarbons were also encountered from the soil samples in many monitoring wells. Simultaneously, Benzene, Toluene, Ethyl-Benzene, Xylenes, were also presented in the AST and flare pit areas.

2.2 Groundwater Transport Model

Developed by Resources & Systems International, Inc., BIOF&T-3D could be utilized to accurately model biodegradation, flow and transport process in the saturated and unsaturated zones in two or three dimensions in heterogeneous, anisotropic porous media or fractured media. BIOF&T's powerful functions allow real world modeling which are not usually available in similar packages. It could model convection, dispersion, diffusion, adsorption, desorption, and

CDI = CW × F

where, CDI = contaminant concentration (mg/l); CW = contaminant concentration (mg/l); F = frequency of exposure during averaging time (min/day)

The results from this study will aid in the risk characterization, which is necessary for the development of a site-specific management plan.
microbial processes based on oxygen-limited, anaerobic, first-order, or Monod-type biodegradation kinetics as well as anaerobic of first-order sequential degradation involving multiple daughter species.

Here, four Light Non-Aqueous Phase Liquid (LNAPL) Benzene, Toluene, Ethylbenene, and Xylene (BTEX) are identified as main contaminants of concern that are considered to contribute most significantly to risks for the exposed populations due to their relatively high concentration, and their dangerous toxicity in nature.

Depending on the inputs of required model parameters, sets of time series contaminant concentration results will be derived. They will be sequentially employed for the risk assessment process.

2.2 Individual Based Probabilistic and Possibilistic Risk Assessment

Uncertainties inherently exist in the environmental process due to sparse and imprecise nature of the available information. Variabilities existing in the individuals make it even more complicated to evaluate the relevant human health risk. Truly, exposures will vary as individual vary in terms of age, gender, health status, exposure duration, propensity to drink water, body size and so on.

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. A number of authors have adapted statistical data related to variables such as drinking water ingestion, soil ingestion, and residential tenure into forms designed for use in probabilistic risk assessment. U.S. Environmental Protection Agency (USEPA) also accepted probabilistic risk analysis and simulation modeling as appropriate tools for risk analysis as CERCLA or "superfund" sites.

2.2.1 Risk Characterization

Contaminant concentrations for the site health risk assessment were gathered during a field effort which included sampling of surface and subsurface of the site and the time series simulation results from groundwater transport model. Exposure pathways and exposure populations were identified based on the observed spatial patterns of site contamination and on-site specific land use and behavior of potential receptors. Here, one potential exposure pathway, groundwater ingestion is carried through the quantitative risk analysis process.

Standard exposure pathway intake models were used to estimate chronic contaminants ingestions.

$$\text{CDI} = \frac{\text{CW} \times \text{IR} \times \text{CF} \times \text{F} \times \text{EF} \times \text{ED}}{\text{BW} \times \text{AT}}$$  

(1)

where, CDI = chronic daily intake of contaminant (mg/kg/day); CW = contaminant concentrations in groundwater (µg/L); IR = groundwater ingestion rate (L/day); CF = conversion factor (10⁶ mg/g); F = relative fraction absorbed from water (100%); EF = frequency of site contact (350 days/365 days); ED = exposure duration (30 years); BW = body weight (kg); AT = averaging time (years)

The results from equation (1) can then be taken as inputs to the risk characterization process. USEPA method for health risk effect is employed and given in the following equation.

Excess Lifetime Cancer Risk (ELCR) = Chronic daily intake (CDI) * Cancer Potenstial Factor (CSF)

(2)

The Slope factor is generally derived from the animal experimental data, and used with administered doses to estimate probability of increased cancer incidence over a life time.

Monte Carlo Simulation

Monte Carlo simulation addresses the weakness of the current risk assessment methods pose in the uncertainty analysis. In extending the regular methods for public health risk assessment, Monte Carlo simulation involves several valuable technologies to estimate both point values and full distributions for the exposure and risk (Burmaster et al., 1989). These extended techniques make the analyses more informative to risk managers and members of the public by giving some perspective of the uncertainty behind the point estimate (Finkel, 1990).

The first step in the Monte Carlo simulation is to determine PDFs to describe sensitive variables in the uncertainty analysis, probability distributions for each of exposure parameter values were estimated using site-specific data, estimates from the literature review, and judgment from receptor behavior. For each uncertain parameter, a single value is generated based on the probability distribution function. These single values are then used in the subsequent model calculations to produce a single answer. A large number of iterations are performed, repeating this procedure of generating random values and performing model calculation. Therefore, the results will finally lead to a distribution of possible values of daily intake rather than a single value (Burmaster et al., 1988). Chronic contamination ingestion values were calculated for each pathway and exposed population using Monte Carlo simulation model as describe below. And Results are then used as inputs for the further risk analysis.

Parameters distribution function and sampling monitoring well data are listed in following table:

<table>
<thead>
<tr>
<th>Table 2: Parameter Distribution Summary for various groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distributioen</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Groundwater ingestion rate (L/day)</td>
</tr>
<tr>
<td>Age group 1: 0-2 years old</td>
</tr>
<tr>
<td>Age group 3: 7-12 years old</td>
</tr>
<tr>
<td>Age group 4: 10-18 years old</td>
</tr>
<tr>
<td>Body Weight (kg)</td>
</tr>
</tbody>
</table>

TIB
Table 1. Concentration of Contaminants in the monitoring well

<table>
<thead>
<tr>
<th>Contaminant Name</th>
<th>Concentration (mg/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzene</td>
<td>0.31</td>
</tr>
<tr>
<td>Ethyl-Benzene</td>
<td>0.00004</td>
</tr>
<tr>
<td>Toluene</td>
<td>0.1</td>
</tr>
<tr>
<td>Xylenes</td>
<td>0.039</td>
</tr>
</tbody>
</table>

Fuzzy Risk Assessment

In the majority of the past and existing risk assessment projects, Federal and statewide environmental guidelines are widely used as the basis of environmental quality evaluation criteria. The Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) authorized the federal government to respond directly to releases of hazardous substances that may endanger public health, welfare and the environment. In addition, The National Oil and Hazardous Substances Pollution Contingency Plan establishes a framework for implementing CERCLA by curbing the process for developing and evaluating appropriate response action for Superfund Sites. And USEPA has also developed risk assessment procedures to address the public health concerns and to ensure that Superfund response actions limit the concentration of hazardous substances in the environment to avoid unacceptable risks to human health (USEPA, 1986). In practical site management, however, these guidelines are mostly conservative and sometimes impractical.

Normally, slope factor, is a chemical-specific constant that describes the carcinogenicity of a compound. It generally derived from animal experiments data and is widely applied to estimate human risk accused by the toxic contaminants (USEPA, 1989). This interspecies conversion factor is a simple number based on the plausible assumption that these toxicities are as harmful to the human being as what they applied to the animals. Because of the assumptions made and methodology used in its derivation, SF values estimated are inherently uncertain. Obviously, serious uncertainty problems exist between the contaminant exposure and potentiality of causing human health cancer risk. A quantitative method, therefore, is required to be employed.

(1) Fuzzy Set Theory

Fuzzy set were first introduced in 1965 by Lotfi Zadeh (1965) to describe imprecisely defined classes or sets that play an important role in human thought processes and communication. In essence, the theory of fuzzy sets is aimed at the development of a body of concepts and techniques for dealing with sources of uncertainty or imprecision that are non-statistical in nature. In classical set theory, an object either belongs to a set or does not, whereas fuzzy set theory allows an object to have partial membership of a set. Using fuzzy sets, it is possible to represent a set A by a membership function value.

The theory of fuzzy sets deals with sets in a universe of discourse U. A fuzzy set X ∈ U is a generalization of the concept of an ordinary set; it is being defined by membership function

\[ x \in \mu_X(x), x \in U \]

(3)

\( X \) is a particular value of \( X \). \( \mu_X(x) \) represents a membership function of \( X \). Interval [0, 1]. The closer \( \mu_X(x) \) is to 1, the more "certain" one is about the value of \( x \).

Fig. 1 (a) and (b) Types of fuzzy membership functions

Fig. 1 presents two types of fuzzy membership functions (triangular and trapezoidal) for illustrating uncertainties associated with the parameter. For example, The angular membership function means that (i) X is most likely equal to C, and (ii) values lower than a or greater than b are considered impossible for X. A membership function is normally defined based on characteristics of the uncertain information.

Fuzzy logic can be considered as a generalization of the Boolean logic, by extended Boolean logic to handle the notion of partial-truth–truth-values between and including 'completely true' and 'completely false'. Fuzzy logic uses a soft linguistic type of variables, which are defined by continuous range of truth-values or fuzzy membership functions in the interval [0, 1] instead of the strict binary (T or F) decisions and assignments. It is the best tool to analyses and simplifies data, which characterized by vague conception or are subjective by incorporation of fuzzy sets (Zadeh, 1965).

(2) Fuzzy Relations

Fuzzy relationships between fuzzy variables defined on different universes of discourse through the fuzzy conditional statement or linguistic implication

\[ X = Y \text{ or } \mu_X(u) \neq \mu_Y(v) \]

Which links the conditional or antecedent set X defined by \( \mu_X(u) \neq \mu_Y(v) \) with the consequence or output set Y defined by \( \mu_Y(v) \). Let

\[ R = \{(x, y) | \mu_X(x) \leq \mu_Y(y) \} \]

as the Cartesian product of \( X \times Y \), is called a fuzzy relation on \( X \times Y \).

Then an m-dimensional fuzzy vector \( A \) can be obtained as:

\[ A = (a_1, a_2, ..., a_m) \]

According to the principle of fuzzy set operation, \( B \) can be determined by a max-min or max*-composition (Zimmermann, 1985). For the max-min composition,

\[ b_{ij} = \sum_{k=1}^{n} a_{ik} \cap b_{kj} \]

Then an m-dimensional fuzzy vector \( B \) can be obtained as:

\[ B = A \times R = (b_1, b_2, ..., b_m) \]

(4)

For the max*-composition,

\[ b_{ij} = \bigcup_{k=1}^{n} a_{ik} \bigvee b_{kj} \]

(5)

(3) Fuzzy Risk Analysis

Two fuzzy sets U for toxic contaminants and V for different age group are initially defined as following:
\[ U = \{ u_i \} \text{ where } \forall i \]

\[ V = \{ v_j \} \text{ where } \forall j \]

where \( u_i \) represents contaminant \( i \), and \( v_j \) is for age group \( j \).

A fuzzy subset of \( U \times V \), which is a binary fuzzy relation from \( U \) to \( V \), can be characterized through the following membership function:

\[ R : U \times V \rightarrow [0,1] \]

Thus, we have fuzzy relation matrix:

\[ R = [ r_{ij} ] \text{ where } x_{ij} = \mu_{ij}(x) \text{ is the membership function of different contaminant intake to cancer risk for contaminant } i \text{ versus different age group } j \text{ which is a function of contaminant concentration and age group.} \]

The membership function of fuzzy set \( A \) to input factor \( x \), denoted as \( \mu_{A}(x) \{ 0 < \mu_{A}(x) < 1 \} \), will be defined as a linear function (Figure 2). It requires two parameters \( (a_i, b_i) \) which denoted as the lower and upper bounds of allowable exposure dose. The relevant standards for setting \( a_i \) and \( b_i \) are considered and acquired from the orientation materials and established standards from USEPA and published statewide and local government standards. For example, USEPA and California local environmental department respectively take 0.005mg/L and 0.001 mg/L as federal and local guidelines for Benzene concentration in the groundwater to assess the potential health carcinogenicity. And Maximum Contaminant Level Goals (MCLG) for benzene has been set at zero because USEPA believes this level of protection would not cause any of the health effects.

For age group 5, referring to above-mentioned guidelines, the most conservative standard 0.0mg/L and most popular standard 0.005mg/L are used to estimate \( a_i \) and \( b_i \). Taking those two values as concentration inputs into the contaminant intake calculation equation, the corresponding ingestion dose-based standards will be derived. Following the similar concentration standard acquisition methodology, the respective concentration standards and ingestion dose-based standards for the rest of three contaminants can be obtained.

Instead of having a detailed distribution for explaining the uncertainties, we have the membership grade of \( x_{ij} \) (Intake Dose) to cancer risk calculate as following:

\[ \mu_{ij}(x) = \frac{1}{1 + e^{-(x - a_i)/(b_i - a_i)}} \]

Fig. 2 relationship between contaminant intake and cancer possibility

\[ r_{ij} = \begin{cases} 0 & \text{if } x_{ij} < a_i \\ (x_{ij} - a_i)(b_i - a_i)/(b_i - x_{ij}) & \text{if } a_i < x_{ij} < b_i \\ 1 & \text{if } x_{ij} > b_i \end{cases} \]

Thus, we got the matrix format for equation (10).

Since individual contaminant will have different overall impacts on human health, we can construct a weighting set on \( U \) as:

\[ W = (W1, W2, W3, W4) \text{ and } \sum W_i = 1 \]

\[ W = R \leftrightarrow \mu_{w,R} \]

here we define \( \cdot \) as a max-\( \cdot \) composition, so the membership grade

\[ \mu_{W,R} = \sum (\mu_{w} \cdot \mu_{R}) \]

and the result

\[ B = (0.81063, 0.8909, 0.8048, 0.724268, 0.723402) \]

Which could be explained as the possibility that Age1, Age2, Age3, Age4, Age5 could have when exposed to the sample contaminant concentration. Therefore, besides knowing only the membership function value of cancer risk for each contaminant for different age group, and integrated possibility of causing cancer risk by concerned contaminants for different group is also presented as well.

2.3 Extended Health Risk Assessment Approach by Using Fuzzy Set Theory

The limitation of conventional overall risk assessment calculation method, which was spreadly accepted, occurs in the procedure of final overall risk evaluation. Regardless of the physical and chemical characteristic differences or interrelationships for both carcinogens and non-carcinogens, the overall cancer risk or non-cancer risk measurement for exposure to multiple carcinogens and non-carcinogens is normally the risk summarization of individual contaminant to provide the final measurement of risk.

We introduced the concept of constructing risk analysis by fuzzy set theory in the above measurement of the individual health risk assessment. However, it's by the number term of
the occurrence possibility, which is difficult to communicate with the non-expert. In this part, a methodology for evaluating risk assessment by linguistic value and fuzzy relation analysis is utilized subsequently. As described previously, the use of fuzzy sets will allow an analyst to communicate degree of health risk of individual elements to people in a readily understandable language. Once these individual risk elements are communicated, fuzzy set theory would then permit an evaluation of the risk of human health to contaminated waste in linguistic variable.

Based on the previously established criteria, consideration of public health, federal and state wide regulatory limits for groundwater, and well accepted drinking water quality standards, making use of the survey results from experts in the environmental field, sets of linguistic value-supported and more detailed risk level criteria according to the allowed ingestion dose were established. Here, our study is mainly focused on the age group 5.

Table 3. Criteria of Risk Levels Under Different Concentrations for each contaminant (μg/L)

<table>
<thead>
<tr>
<th></th>
<th>Clean</th>
<th>Slightly Contaminated</th>
<th>Contaminated</th>
<th>Significantly Contaminated</th>
<th>Extremely Contaminated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzene</td>
<td>0</td>
<td>1.32</td>
<td>2.55</td>
<td>3.77</td>
<td>5.5</td>
</tr>
<tr>
<td>Ethylbenzene</td>
<td>13.2</td>
<td>174.9</td>
<td>190.6</td>
<td>247.7</td>
<td>347.7</td>
</tr>
<tr>
<td>Toluene</td>
<td>40</td>
<td>254</td>
<td>250</td>
<td>360</td>
<td>260</td>
</tr>
<tr>
<td>Xylenes</td>
<td>8.8</td>
<td>186.6</td>
<td>204.4</td>
<td>332.2</td>
<td>50000</td>
</tr>
</tbody>
</table>

Thus, we have fuzzy relation matrix:

\[
R = \{ r_{ij} \mid i = 1, 2, 3, 4; j = 1, 2, 3, 4, 5 \} \quad (18)
\]

where \( r_{ij} \) is the membership function of contaminant \( i \) versus different cancer risk level \( j \), which is a function of contaminant concentration and risk level criteria.

The groundwater monitoring data for four Volatile Organic Petroleum Contaminants (VOPC) can be presented as \( C_i \).

The membership grade between \( C_i \) and the polluted level grade \( j \) can be calculated, according to \( v_i \), the criterion for pollutant / at polluted level \( j \).

1. When \( C_i \in [v_{ij-1}, v_{ij}] \)

\[
r_{ij} = \frac{\log_{10} c_i - \log_{10} v_{ij-1}}{\log_{10} v_{ij} - \log_{10} v_{ij-1}}
\]

(19)

2. When \( C_i \in [v_{ij}, v_{ij+1}] \)

\[
r_{ij} = \frac{\log_{10} v_{ij} - \log_{10} c_i}{\log_{10} v_{ij+1} - \log_{10} v_{ij}}
\]

(20)

3. When \( C_i \leq v_{ij} \) or \( C_i \geq v_{ij+1} \)

\[
r_{ij} = 0, \quad \forall i, j
\]

(21)

Thus a parameter of fuzzy related matrix can be obtained as follows:

\[
R^2 = \left[ \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij} \right]_{i,j=1,2,3,4}^{n=1,2,3,4}
\]

Similarly, a weighting set for petroleum contaminants and the max^* composition methodology of set \( W \) and \( R \) defined previously are utilized

For this case, \( B = 0.03654, 0.05364, 0, 0, 0.71 \)

And MaxB = 0.71 (Extremely Contaminated)

Therefore, we can determine the risk level for the sampling well is extremely contaminated, high demands exist for the future remediation.

3. GIS SUPPORTED RISK ASSESSMENT

Integrating GIS with environmental risk models can provide a more meaningful interpretation of the problem within a georeferenced environment. A few studies have focused on using GPS for environmental risk assessment. Chen et al. (1996) provided linking GIS with a groundwater model and decision support system for the purpose of a petroleum contaminated site assessment. Miller et al. (1996) conducted a project using GIS to calculate human health risks at a large military facility. For example, spatial and temporal attributes of
contaminants risk can be mapped in ArcInfo coverages. Each coverage is linked to an attribute table so that information is available on the individual features, or records, of the theme. Maps can be created to convey geographic features and relationships along with the results of a data analysis. Besides data-mapping advantage, GIS also has the ability to query specific information in a theme and dynamically link external database.

The general GIS-supported framework is showed in Fig. 4.

![GIS User Interface Diagram](image)

### 3.1 The Digital Base Design

It consists of two major components: a spatial database and a tabular database. The spatial database contains the GIS shape files and coverages of geographic features related to regional and facility information. These features are topologically referenced to the real world. The regional information required by spatial database can be obtained from various source, it's always a time consuming and tedious search work. Most of these sources contain ArcInfo coverages in export format (*.e00), which can be imported into ArcView using build-in ImportPT command.

Obstacles exist when big-scale coverages are not available from Internet sources, federal or local department. Fortunately, options available for building a facility database. One method is first through aerial mapping to produce orthophotos, and then digitize features from the orthophotos. It is accurate but very expensive. Digitizing the facility features using existing maps or photographs is the alternative that is easy-taken and was employed in the research process. Through accessing to a digitizing tablet, layering coverage can be accomplished directly with the Digitizer extension in ArcView (Fig. 5).

![Digitizing Example](image)

Tabular database contains information about environmental measurements at the site, such as contaminants concentration data, groundwater levels, sampling position, soil profiles and so on. An efficient form of organizing data is through the use of a relational database. The database is stored in one file, but this file can contain multiple tables, all of which relate to the particular subject but contain different data related to that subject. However, the current capabilities of ArcView do not provide an efficient way to link multiple tables since each feature can only have one associated record. Under such circumstance, Microsoft Access was used to create the relational database. Interrelationship within tables was established to link each other, which could be used to execute query function to create new relationships (Fig. 6).

![Relational Database](image)

### 3.2 Customized Graphic User Interface

ArcView's Advanced Customized Graphic User Interface operation allows the user to create self-demand user interface with more flexible and realizable function. Through executing easy event-driven Avenue script "System" and "External Script" type, the groundwater transport model simulator BIOFAT could be called externally by ArcView. The results from BIOFAT simulator can be compiled and saved as Text Delimited Format, which can be securely import into MS Excel to proceed the risk analysis using proposed environmental risk assessment methodology.

### 3.3 Connecting Spatial and Tabular Database and Spreadsheet

The predicted concentration data resulting from the groundwater simulator and sampling data were imported to Excel spreadsheet to process the calculation step of environmental risk analysis. When the calculation was done, the risk results were connected with MS Access through the "Get External Data" function. Advanced ArcView Database dynamically link function allows the user to import selected sets of data from another database utility into GIS using an Open Database Connectivity (ODBC) driver and build-in data management system which use Structured Query Language (SQL). The imported data can then be joined to an existing coverage or used to create a new coverage with Avenue scripts or Add Event Theme command in ArcView.

The Spatial Analyst extension of ArcView was used to interpolate grid surface and contours between point measurements. For each figure, the contours of risk level for individual risk and overall risk level were created using Inverse Distance Weighting (IDW) interpolation scheme, which interpolates a values for each point in the study area using the inverse distance squared between the location of the current point and the measurement location as a weighting parameter (Fig. 7).
4. CONCLUSIONS

In this paper, a GIS supported environmental risk assessment using stochastic and fuzzy set theory has been proposed for the petroleum-contaminated site. Employing the general accepted regulation by USEPA, federal or local government, and utilizing the expert opinions on the risk standard through question survey. Environmental risk calculation in occurrence possibility term for different age group and linguistic term risk assessment have performed and the reasonable risk results produced.

Integration of GIS technology with environmental risk assessment has shown the capacity to maximize the utilization value of available information. Through the creation of spatial database and facility relational database, the advanced data exchange and query function, and appropriate data which is necessary for contaminant concern, risk-based decision presentation has been determined, and relevant site physical and chemical characteristic information can be presented spatially.

With the further enhancement of spatial and monitoring data quantity and quality, a more comprehensive multi-pathway and multi-scenario case study will be explored in the future research improvement.

REFERENCES