Stochastic Risk Assessment of Groundwater Contamination under Uncertainty: A Canadian Case Study

J. B. Li[1,2*]

[1]Environmental Engineering Program, University of Northern British Columbia, Prince George, British Columbia, Canada
[2]Energy and Environmental Research Center, North China Electric Power University, Beijing 102206, China

ABSTRACT. This study presents a stochastic approach for risk assessment of groundwater contamination through incorporating a stochastic subsurface contaminant transport and fate modeling system within a general risk assessment framework. The uncertainties associated with soil properties (e.g. soil porosity and permeability) were addressed through the probabilistic method, and the resulting uncertainties in risks of groundwater contamination were then identified under different remediation scenarios. The method was applied to a petroleum-contaminated site in western Canada. Three remediation scenarios with different clean-up efficiencies (0, 60, and 90%) were examined for a planning period of 10 years to study the risks introduced by xylene. The concept of hazard index was applied to characterize the associated non-carcinogenic risk. Three risky zones with different mean hazard index distributions were identified, and the probabilities of violating assumed hazard index criteria at the study site were also presented. The obtained results indicated that the proposed method could effectively identify risky zones with different risk levels and violation probabilities under various remediation actions, and they are directly useful for the decision maker to gain insight into the study site and to make remediation decisions.

Keywords: Contamination, groundwater, modeling, risk assessment, stochastic, uncertainty

1. Introduction

Effective management of contaminated soil and groundwater systems has been a principle task of environmental professionals and regulatory agencies (USEPA, 2001; Khadam and Kaluarachchi, 2003). Among various tasks of contaminated site management, risk assessment is a crucial component. It offers a means for decision support through carefully assessing and ranking severity of site contamination, and thus allowing identification of critical issues for mitigation actions (Carrington and Bolger, 1998; Li et al., 2006). Risk assessment is usually based on the support of mathematical models by projecting subsurface contaminant behavior into the future (Chen et al., 2000), while the efficiency of the mathematical modeling efforts usually requires sufficient knowledge of the subsurface hydrogeological conditions throughout a contaminated site. However, this knowledge is often limited by various uncertainties associated with soil and contaminant properties, and the risk is thus inherently linked with uncertainties (Hoffman et al., 1999). It is recognized that the success of contaminated site risk assessment depends significantly on whether the contaminant transport and fate models have appropriately quantified and incorporated the related uncertainties into the simulation processes (Maxwell, 1998; Li et al., 2003).

During the past years, a significant number of literatures

and regulatory materials have been published on methods for conducting uncertainty-based environmental risk assessment at contaminated sites. For example, Goodrich and McCord (1995) applied Monte Carlo methods to account for parameter uncertainties associated with ground-water flow and solute transport for supporting exposure assessment. Piver et al. (1998) developed a method to evaluate risks from contaminated groundwater by solving a stochastic one-dimensional contaminant transport model through the Monte Carlo simulation technique. Lo et al. (1999) applied Monte Carlo simulation technique to predict the probability for a contaminant to exceed its maximum allowable concentration in groundwater underneath a landfill site, based on a one-dimensional solute transport model. Lee et al. (2002) applied the USEPA guidelines and the empirically measured contaminant levels and exposure parameters to estimate health risk in a groundwater-contaminated community after on-site remediation, by using Monte Carlo simulations to account for uncertainties. Li et al. (2006, 2007) proposed a fuzzy-stochastic risk assessment approach for examining uncertainties associated with both source/media conditions and evaluation criteria in the contaminated site management system. However, many of the previous risk assessment studies were either based on simple numerical models or for short-term only. Such simplification cannot sufficiently characterize the complexities especially existing at petroleum-contaminated sites. In addition, there are only few studies on integrating advanced numerical subsurface models, such as multi-phase multi-component contaminant transport and fate models, within an uncertainty-based

---

* Corresponding author: li@unbc.ca

risk assessment framework (Li et al., 2007). This kind of integration, however, can effectively facilitate the assessment of long-term risks at hydrocarbon-contaminated sites. Such a long-term risk assessment will then be crucial for providing more realistic decision support for pollution mitigation strategies in terms of effective site management, and bring enormous environmental and economic benefits.

As an extension of the previous efforts, the objective of this study is to present an integrated approach for environmental risk assessment of groundwater contamination, through incorporation of a stochastic multiphase compositional modeling system within a general risk assessment framework. The method will be applied to a petroleum-contaminated site in western Canada. Three remediation scenarios with different efficiencies (0, 60 and 90%) will be examined for a planning period of 10 years at the study site. Xylene is selected as a constituent of interest in hydrocarbon, and the concept of hazard index is applied to study the resulting non-carcinogenic risk. Three risky zones with different mean hazard index distributions will then be identified based on local environmental guideline and assumed hazard index criteria, and the probabilities of violating the hazard index criteria will also be presented. The obtained results will be directly useful for the decision maker to gain insight into the study site and to make remediation decisions.

2. Methodology

2.1. Monte Carlo Simulation of Contaminant Transport

Modeling for pollutant transport in subsurface requires inputs of various physical, chemical and biological parameters. However, some fundamental parameters, such as soil permeability and porosity, are generally difficult to acquire with accurate and deterministic values, and a number of uncertainties are associated with them (Labieniec et al., 1997; Schnabel et al., 2004). Many researchers have used statistical or stochastic techniques to quantify these uncertainties (Hoffman et al., 1999, Hu and Huang, 2002). Among the stochastic techniques, the most popular approach so far has been the Monte Carlo simulation. Monte Carlo techniques utilize repeated executions of numerical models to simulate stochastic processes. Each execution of the model produces a sample output. The output samples can then be examined statistically and distributions can be determined (USEPA, 1997). Monte Carlo techniques have a number of advantages, such as: (a) the ability to handle uncertainty and variability associated with modeling parameters; (b) the capability to be applied in deterministic modeling framework; and (c) the flexibility of dealing with various types of probability distributions for the model inputs (Maqsood et al., 2003).

The Monte Carlo simulation algorithm has been developed and incorporated within a multiphase multi-component simulator named UTCHEM (UTA, 2000; Li et al., 2003) in this study, and a stochastic subsurface modeling system was then established. The related procedures can be summarized as follows (Li et al., 2007): (a) generate random numbers

between zero and one for each stochastic input parameter; (b) transform the random numbers to the corresponding random variates based on the specified probabilistic distribution for each parameter; (c) store the generated random variates into an array for each parameter; (d) obtain a value from the array for each parameter and set it as the deterministic input parameter in the multiphase multi-component model; (e) compute the contaminant concentrations with the multiphase compositional numerical model for each Monte Carlo run; (f) store the resulting outputs of the contaminant concentrations for each Monte Carlo run for further statistical analysis; (g) repeat steps (a) to (f) for a number of times (specified number of Monte Carlo runs); (h) stop computation of contaminant concentrations when the specified number of Monte Carlo runs has been done, and exit to step (i), and (i) analyze the contaminant concentrations and compute the statistic moments (i.e., peak concentration at the domain for each run, mean and standard deviation of the concentrations over the specified number of times of Monte Carlo runs).

2.2. Approaches for Risk Assessment

The related risk characterization associated with contaminated site can usually be conducted through environmental-guideline-based risk assessment (ERA) and health risk assessment (HRA) (Carrington and Bolger, 1998). The ERA approach is to compare contaminant concentration with the corresponding groundwater quality guideline and to characterize the risk due to guideline violation. The HRA is to identify the risk of health impact due to chronic intake of contaminant. To quantify human health risks, pollutants are characterized as carcinogens and non-carcinogens. HRA includes excess lifetime cancer risk (ELCR) model for assessing carcinogenic risk, and hazard index (*HI*) model for non-carcinogenic risk. In this study, xylene in hydrocarbon is selected as a constituent of interest to examine the related risk. As a result, the concept of *HI* is applied. The degree of exposure to a chemical is a function of many variables described as follows (USEPA, 1992):

$$CDI = CW \cdot IR \cdot EF \cdot ED / (AT \cdot BW) \tag{1}$$

where *CDI* is the chronic daily intake (mg/kg·d), *CW* is pollutant concentration in groundwater (mg/L), *IR* is human ingestion rate (L/day), *EF* is exposure frequency (days/year), *ED* is average exposure duration (year), *BW* is average body weight (kg), and *AT* is averaging time ($AT = 365 \cdot ED$ days). The *HI* can then be calculated by the following equations:

$$HI = CDI / RfD \tag{2}$$

where *RfD* is a reference dose (mg/kg·day). The calculated hazard index (*HI*) indicates the extent of violating a health risk limit (e.g. *RfD*) of the non-carcinogenic contaminant, and usually the acceptable level of risk denoted by *HI* is defined as lower than 1.0.

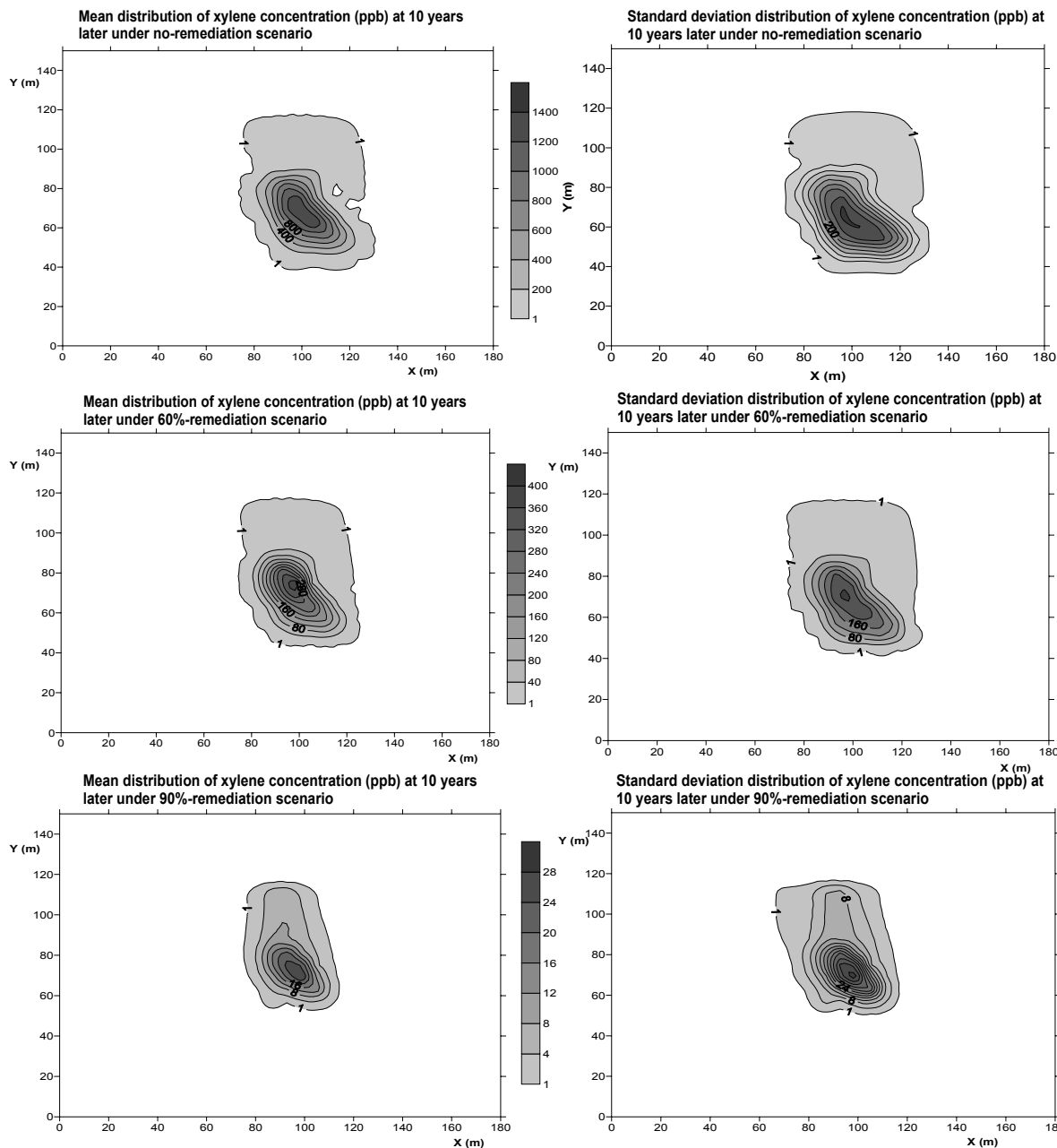The ERA and HRA may obtain different insights about

Figure 1. Mean and standard deviation distributions of xylene concentration (ppb) at 10 years later under different remediation scenarios.

risks. For example, the Canadian guideline of xylene is 0.3 mg/L when the groundwater is used for drinking water supply (SERM, 2002), and any concentration greater than 0.3 mg/L will be regarded as risky. However, 0.3 mg/L of xylene concentration in groundwater is corresponding to a *HI* of 0.041 (when using parameter values of *IR* = 2 L/day, *EF* = 350 days/year, *ED* = 30 years, *BW* = 70 kg) by applying the *RfD* value of 0.2 mg/kg·d as recommended by USEPA (2003), which is considered to be an acceptable level of risk since the *HI* is lower than 1.0. In order to combine the concepts of both ap-

proaches, the risk is characterized into the following three levels in this study: (a) slightly risky with *HI* in the range of 0.003 to 0.041, where *HI* of 0.003 is corresponding to a xylene concentration of 0.02 mg/L which is the strictest guideline (Li et al., 2007); (b) risky with *HI* in the range of 0.041 to 1.0, and (c) highly risky with *HI* greater than 1.0. The contaminant concentration denoted as *CW* in Equation (1) will be a stochastic variable due to the uncertainties in the input parameters of the multiphase compositional simulator, leading to uncertainties in the calculated *HI* as described by Equation

Hazard index (HI) distribution 10 years later under no-remediation scenario

Slightly risky zone (with HI in the range of 0.003 to 0.041)
Risky zone (with HI in the range of 0.041 to 1.0)
Highly risky zone (with HI in the range of greater than 1.0)

Hazard index (HI) distribution 10 years later under 60%-remediation scenario

Slightly risky zone (with HI in the range of 0.003 to 0.041)
Risky zone (with HI in the range of 0.041 to 1.0)
Highly risky zone (with HI in the range of greater than 1.0)

Hazard index (HI) distribution 10 years later under 90%-remediation scenario

Slightly risky zone (with HI in the range of 0.003 to 0.041)
Risky zone (with HI in the range of 0.041 to 1.0)
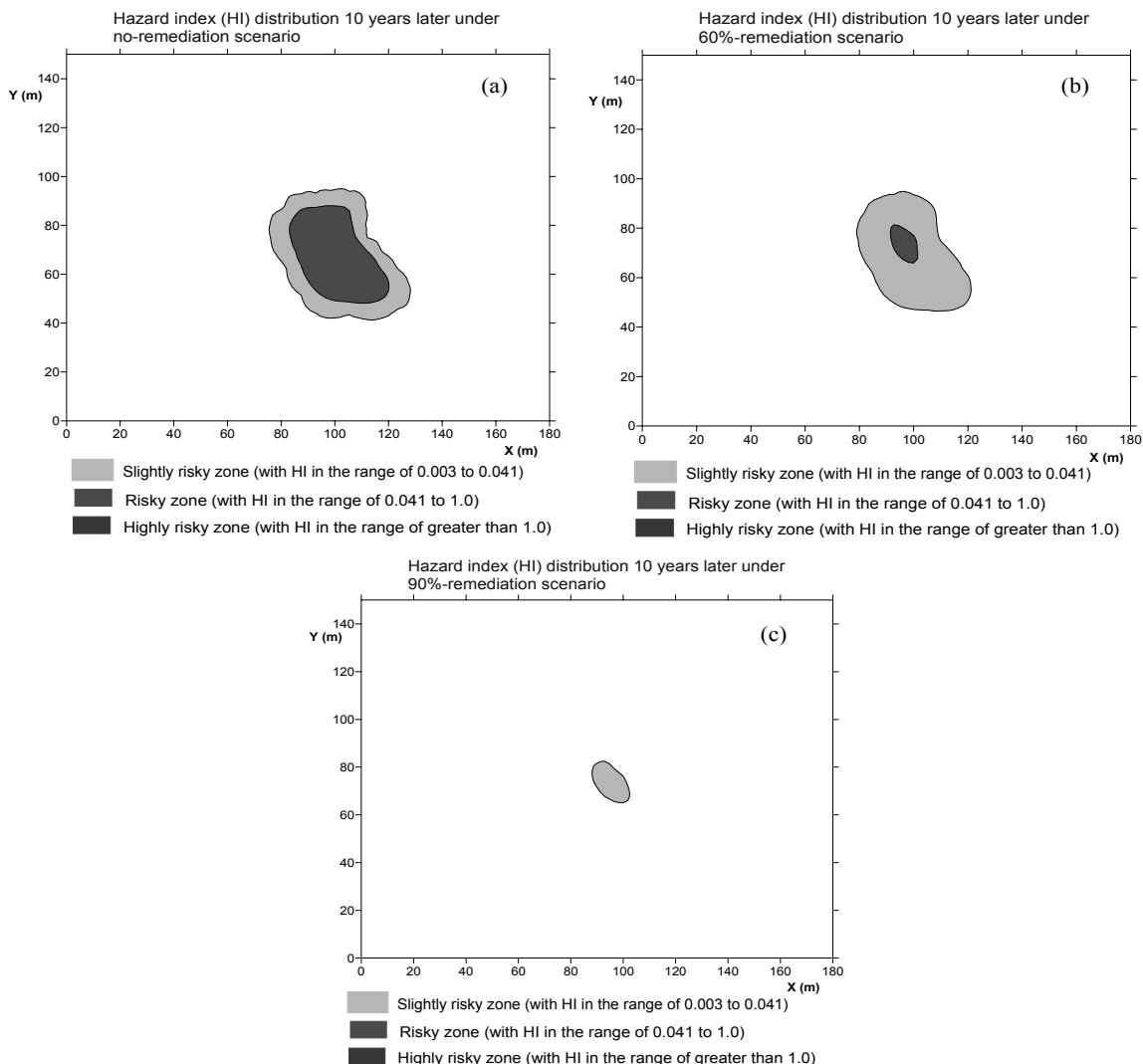Highly risky zone (with HI in the range of greater than 1.0)

Figure 2. Distribution of different risky zones at 10 years later under various remediation scenarios.

(2). As a result, through Monte Carlo simulation, the probability (P_F) under which the hazard index (*HI*) exceeds the evaluation criteria can be described as follows:

$$P_F = P(HI > HI_s) = 1 - F(HI_s) \tag{3}$$

where *HI* is the hazard index, HI_s is the hazard index criteria for risk evaluation (assumed as 0.041 in this study), and $F(HI_s)$ is the cumulative distribution function (CDF) of hazard index which can be obtained from Monte Carlo simulation results.

3. Case Study

A petroleum-contaminated site in western Canada was selected as a case study site to illustrate the proposed risk assessment approach. The site has been operated as a natural gas processing plant to remove naphtha condensate from the na-

tural gas stream prior to transport to a regional transmission line (Li et al., 2007). Throughout the history of the site, naphtha condensate, a waste liquid removed from the gas by a series of scrubbers, was disposed of in a previous underground storage tank (UST). Due to leakage of this UST in the past years, the site has been contaminated. During the past several years, a number of field investigations have been conducted to examine the site hydro-geologic conditions and related pollution problems. The groundwater flow is towards the northwest direction at the site, which was obtained through multiple site investigation results. Various forms of the pollutants were found at the downgradient directions, and they include free product, vapor phase hydrocarbon, dissolved hydrocarbon in groundwater, and trapped hydrocarbon in the soil. The overview of the site can be found in Li et al. (2007). In this study, the xylene from petroleum contaminants is selected as the constituent of interest. The highest xylene concentration of 1.12 mg/L was observed at a monitoring well located in the

west of the previous UST, and this concentration is higher than the regulated local environmental guideline (e.g. 0.3 mg/L). As a result, the groundwater contamination at this site may pose serious health risk to the surrounding communities, and conducting effective risk assessment is desired.

4. Result Analysis

4.1. Monte Carlo Simulation of Xylene Transport

The developed stochastic modeling system was used for simulating xylene transport and fate in the groundwater. The detailed descriptions of modeling domain, soil properties and modeling input parameters can be found in Li et al. (2007). Among various input parameters, the values of soil permeability (K_{xx}) and porosity (ϕ) are considered to be lognormally and normally distributed, respectively. The mean values and standard deviations of K_{xx} and ϕ for three on-site soil types (sand, clay till, and silty clay) are used for Monte Carlo simulation. The mean and standard deviation distributions of xylene concentration at 10 years later under three remediation scenarios with cleanup efficiency of 0%, 60% and 90% are presented in Figure 1. It is illustrated that the uncertainties in input parameters will result in significant uncertainties in the output concentrations that are calculated by the numerical simulator.

The resulted xylene concentration uncertainties could be characterized by the corresponding standard deviations. For example, under the no-remediation scenario, the maximum mean concentration of xylene in groundwater at the study site 10 years later is predicted to be 1.365 mg/L as shown in Figure 1, while the maximum standard deviation is predicted as 0.331 mg/L; under the 60%-remediation scenario, the maximum mean concentration at 10 years later is predicted as 0.389 mg/L, while the maximum standard deviation is 0.245 mg/L; under the 90%-remediation scenario, the predicted maximum mean concentration at 10 years later is 0.027 mg/L, while the maximum standard deviation at the site after 10 years is 0.046 mg/L. These uncertainties in concentration prediction would result in significant uncertainties in the related risk assessment, and they should be incorporated within the risk assessment process.

4.2. Spatial Distribution of Hazard Index

The proposed stochastic multiphase compositional modeling system facilitates dynamic modeling of pollutant fate and transport under a number of remediation scenarios by considering parameter uncertainties, and the related results provide sound bases for clarifying contamination conditions at the site in the future. The simulated mean xylene concentrations in groundwater and the relevant plume extents at 10 years later under 0, 60 and 90% remediation actions, are used as crucial information for assessing the risks at the site due to groundwater contamination when the on-site groundwater is applied for drinking water supply. The non-carcinogenic risk introduced by xylene is characterized by hazard index (*HI*), and the relevant parameters for risk assessment model (Equa-

tions 1 and 2) include the following values: *IR* = 2 L/day, *EF* = 350 days/year, *ED* = 30 years, *BW* = 70 kg, and *RfD* = 0.2 mg/kg.d. Three different risky zones are characterized based on the assumed *HI* evaluation criteria as described in section 2.2.

Figure 2(a) shows spatial distribution of the predicted mean *HI* at 10 years later under the no-remediation scenario. The results illustrate that the slightly risky zone will extend almost 44 m northwest of the site's previous UST with an area of about 922 m^2; the risky zone will extend over 37 m northwest of the former UST, and the total area of this zone is about 1,020 m^2; there is no highly risky zone which has the *HI* of greater than 1.0. Figure 2(b) presents spatial distribution of predicted mean *HI* 10 years later under the 60%-remediation scenario. The results indicate that the slightly risky zone will extend over 40 m northwest of the previous UST with an area of 1,290 m^2; the risky zone will extend over 29 m northwest of the former UST with an area of 100 m^2, indicating that 60%-remediation efficiency action would significantly reduce the risky zone after 10 years; there is no highly risky zone. Figure 2(c) shows the mean *HI* at 10 years later under the 90%-remediation scenario, indicating the non-existence of risky zone any more and significant shrinkage of the slightly risky zone with an area of around 170 m^2 by using highest cleanup efficiency actions.

4.3. Probability of Violating Risk Evaluation Criteria

The mean of hazard index was used for characterizing three different risky zones based on the assumed *HI* evaluation criteria as described above. The Monte Carlo simulation results can also be presented in the form of probability distribution, and the probability of *HI* violating the related evaluation criteria at the study site would provide more in-depth insights of risks for the decision maker. With a conservative consideration, the predicted peak *HI* for each Monte Carlo simulation run is identified for further risk assessment. The modeling results indicate that at 10 years later, the peak *HI* under the no-remediation scenario is in the range of 0.128 to 0.267, with a mean of 0.185 and standard deviation of 0.043; the peak *HI* under the 60%-remediation scenario ranges from 0.0003 to 0.141, with a mean of 0.069 and standard deviation of 0.049; under the 90%-remediation scenario, the peak *HI* ranges from 0.0 to 0.036, with a mean of 0.005 and standard deviation = 0.011. It is indicated that significant uncertainties exist in the predicted *HI*, and the peak *HI* would be significantly reduced when a higher remediation-efficiency action is taken. For example, for the 90%-remediation scenario, there is no chance of *HI* exceeding 0.041 which is corresponding to the local environmental quality guideline, indicating the non-existence of risky zone at the study site 10 years later under 90%-remediation actions.

The probability of *HI* violating the related evaluation criteria at the study site can be obtained from the cumulative distribution functions (CDF) of peak *HI* as shown in Figure 3. In this study, the *HI* evaluation criteria of 0.041 is used to examine the probability of reaching risky level at the site. Figure 3(a) shows the CDF of peak hazard index at 10 years

later under the no-remediation scenario. It is observed that 0% of *HI* is less than 0.041, indicating 100% of probability of the existence of risky zone at the site at 10 years later under this scenario. Figure 3(b) presents the CDF of peak *HI* at 10 years later under the 60%-remediation scenario, and it is observed that 35% of *HI* is less than 0.041, indicating 65% of probability of the existence of risky zone at the site 10 years later under this scenario. Figure 3(c) shows the CDF of peak *HI* at 10 years later under the 90%-remediation scenario, and it is observed that 100% of *HI* is less than 0.041, indicating 0% of probability of the existence of risky zone at the site at 10 years later under the 90%-remediation scenario.
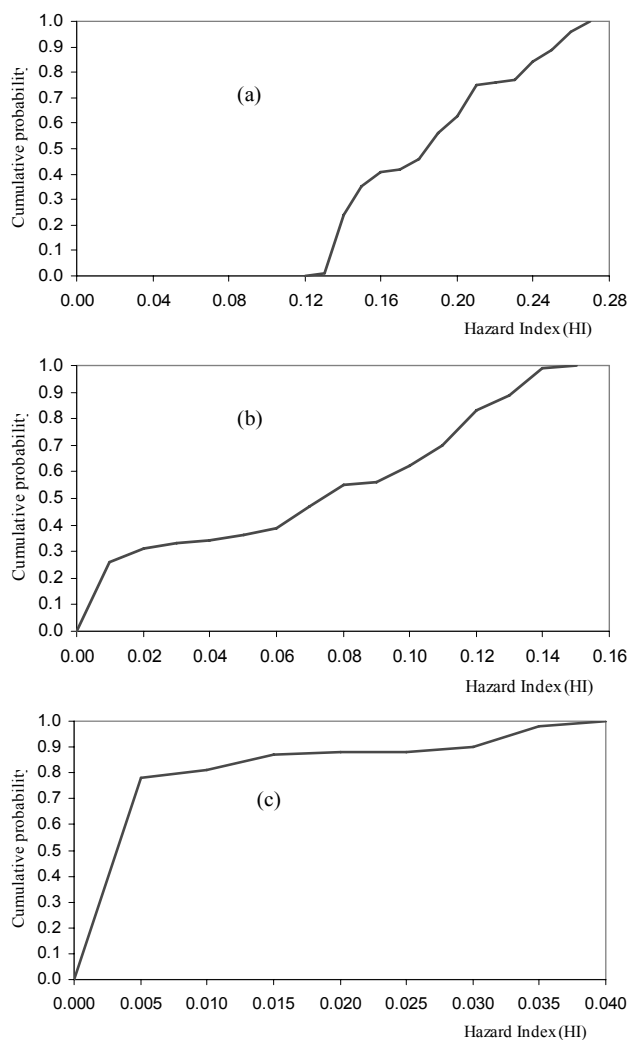

Figure 3. Cumulative distribution functions of hazard index at 10 years later, (a) no-remediation, (b) 60%-remediation, (c) 90%-remediation.

5. Conclusions

In this study, an integrated approach for risk assessment of the subsurface contamination was developed through incorporation of a stochastic multiphase compositional subsurface

modeling system within a general risk assessment framework. This development was based on (a) Monte Carlo simulation for the fate of contaminants (e.g. xylene) in subsurface via a 3-D multiphase compositional numerical model to account for stochastic uncertainties in soil properties (e.g. soil porosity and permeability); (b) examination of simulation results that were expressed as mean and standard deviation; (c) quantification of non-carcinogenic risk using the concept of hazard index (*HI*); (d) characterization of various risky levels based on assumed *HI* evaluation criteria and the mean *HI* distribution, and (e) quantification of the probability of the existence of risky zones based on the obtained cumulative distribution functions of *HI*. The proposed stochastic risk assessment approach was applied to a petroleum-contaminated groundwater system in western Canada. Three remediation scenarios with different efficiencies (0, 60%, and 90%) were examined for a planning period of 10 years at the study site, while xylene was selected as a constituent of interest in hydrocarbon. The simulation results indicated that the uncertainties in soil property parameters would result in significant uncertainties in risk evaluation. Three risky zones with various mean hazard index distributions were identified based on local environmental quality guideline and selected hazard index evaluation criteria when the on-site groundwater was assumed for drinking water supply. The probabilities of violating assumed hazard index criteria (with a *HI* of 0.041) were also presented, and the obtained results would be directly useful for the decision maker to gain in-depth insight into the study site and to make remediation decisions.

References

Carrington, C.D. and Bolger, P.M. (1998). Uncertainty and risk assessment. *Human Ecol. Risk Assess.*, 4(2), 253-257.

Chen, Z., Huang, G.H. and Chakma, A. (2000). Risk assessment of a petroleum-contaminated site through a multi-phase and multi-component modeling approach. *J. Pet. Sci. Eng.*, 26(1), 273-282.

Goodrich, M.T. and McCord, J.T. (1995). Quantification of uncertainty in exposure assessments at hazardous waste sites. *Ground Water*, 33(5), 727-732.

Hoffman, F.O., Chambers, D.B. and Stager, R.H. (1999). Uncertainty is part of making decisions. *Human Ecol. Risk Assess.*, 5(2), 255-261.

Hu, B.X. and Huang, H. (2002). Stochastic analysis of reactive solute transport in heterogeneous, fractured porous media: A dual permeability approach. *Transp. Porous Media*, 48(1), 1-39.

Khadam, I. and Kaluarachchi, J.J. (2003). Applicability of risk-based management and the need for risk-based economic decision analysis at hazardous waste contaminated sites. *Environ. Int.*, 29, 503-519.

Labieniec, P.A., Dzombak, D.A. and Siegrist, R.L. (1997). Evaluation of uncertainty in a site-specific risk assessment. *J. Environ. Eng.*, 123(3), 234-243.

Lee, L.J.H., Chan, C.C., Chung, C.W., Ma, Y.C., Wang, G.S. and Wang, J.D. (2002). Health risk assessment on residents exposed

to chlorinated hydrocarbons contaminated in groundwater of a hazardous waste site. *J. Toxicol. Environ. Health*, Part A, 65(3/4), 219-235.

Li, J.B., Huang, G.H., Chakma, A., Zeng, G.M. and Liu, L. (2003). Integrated fuzzy-stochastic modeling of petroleum contamination in subsurface. *Energy Sources*, 25(6), 547-563.

Li, J.B., Liu, L., Huang, G.H. and Zeng, G.M. (2006). A fuzzy-set approach for addressing uncertainties in risk assessment of hydrocarbon-contaminated site. *Water, Air, Soil Pollut.*, 171(1-4), 5-18.

Li, J.B., Huang, G.H., Zeng, G.M., Maqsood, I. and Huang, Y.F. (2007). An integrated fuzzy-stochastic modeling approach for risk assessment of groundwater contamination. *J. Environ. Manage.*, 82(2), 173-188.

Lo, I.M.C., Law, W.K.W. and Shen, H.M. (1999). Risk assessment using stochastic modeling of pollutant transport in landfill clay liners. *Water Sci. Technol.*, 39(10-11), 337-341.

Maqsood, I., Li, J.B. and Huang, G.H. (2003). Inexact multiphase modeling system for the management of uncertainty in subsurface contamination. *Pract. Periodical Hazard., Toxic, Radioact. Waste Manage.* (ASCE), 7(2), 86-94.

Maxwell, R.M. (1998). *Understanding the Effects of Uncertainty and Variability on Groundwater-Driven Health Risk Assessment*, Ph.D. Dissertation, University of California, Berkeley.

Piver, W.T., Duval, L.A. and Schreifer, J.A. (1998). Evaluating health risks from ground-water contaminants. *J. Environ. Eng.*, 124(5), 475-478.

Schnabel, U., Tietje, O. and Scholz, R.W. (2004). Uncertainty assessment for management of soil contaminants with sparse data. *Environ. Manage.*, 33(6), 911-925.

SERM (Saskatchewan Environment and Resource Management) (2002). *Risk Based Corrective Actions for Petroleum Contaminated Sites*, Province of Saskatchewan, Regina, Saskatchewan, Canada.

USEPA (US Environmental Protection Agency) (1997). *Guiding Principles for Monte Carlo Analysis, USEPA EPA/630/R-97/001*, US Environmental Protection Agency, Risk Assessment Forum, Washington DC, pp. 35.

USEPA (US Environmental Protection Agency) (1992). *Guidelines for Exposure Assessment, USEPA 600Z-92/001*, US Environmental Protection Agency, Risk Assessment Forum, Washington DC, pp. 170.

USEPA (US Environmental Protection Agency) (2001). *Abstracts of Remediation Case Studies: Volume 5*, EPA 542-R-01-008, Washington DC.

USEPA (US Environmental Protection Agency) (2003). Integrated Risk Information System: Xylenes (CASRN 1330-20-7). http://www.epa.gov/iris/subst/0270.htm.

UTA (University of Texas at Austin) (2000). *Technical Documentation for UTCHEM-9.0: A Three-Dimensional Chemical Flood Simulator, Reservoir Engineering Research Program*, Center for Petroleum and Geosystems Engineering, University of Texas at Austin, TX.