Setting information priorities for remediation decisions at a contaminated-groundwater site

Hwong-wen Ma ^{a,*}, Kuen-Yuh Wu ^b, Chung-Da Ton ^a

^a Graduate Institute of Environmental Engineering, National Taiwan University, 71 Chou-Shan Road, Taipei 106, Taiwan, ROC

^b Department of Occupational Safety and Health, China Medical College, No. 91, Hsuesh-Shih Road, Taichung, Taiwan, ROC

Received 17 November 2000; received in revised form 6 February 2001; accepted 6 February 2001

Abstract

Many sites of contamination due to inappropriate disposal of hazardous materials or wastes have been found. These sites have the potential of damaging the environment and human health and thus need to be evaluated as to whether and what actions should be initiated. In the decision on whether a contaminated site should be subject to management, the knowledge concerning important parameters that would influence the decision will be beneficial to planning of data collection to support the decision. This paper presents a case study of contaminated site located in northern Taiwan, where the groundwater is contaminated by chlorinated hydrocarbons including trichloroethylene (TCE) and tetrachloroethylene (PCE). A site-specific multimedia risk assessment is performed to estimate the total risk resulting from the contamination. In addition, Monte Carlo simulation, rank correlation coefficients, and decision criteria are combined to develop a methodology for assessing the important of parameters in terms of their influence on the decision. It is found that TCE concentration, vegetable yield, deposition interception fraction of vegetables, and plant surface loss constant, are the four parameters important to the decision-making of the case problem. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Risk assessment; Risk management; Sensitivity; Uncertainty

1. Introduction

Many sites of contamination due to inappropriate disposal of hazardous materials or wastes have been found. These sites have the potential of damaging the environment and human health, and thus need to be evaluated as to whether and what actions should be taken to control the contamination or impacts. Risk assessment, health risk assessment in particular, is one of the most important tools of providing the relevant evaluation. Its role in supporting environmental decision-making is increasingly important because of its

ability of systematically quantifying the health impacts as a result of environmental problems, by logically connecting the source of the problems and the receptors of the impact (US EPA, 1989). For example, a survey indicated that 37 states in the United States are using a risk-based corrective action approach to establish cleanup levels (Barken et al., 1996; Lahkim and Garcia, 1999).

Risk assessment entails evaluations of the source, transport, fate, and exposure of chemicals of concern. Increasingly, risk assessment methods are moving from generic or regional assessment to site-specific assessment, because the health risk received by a receptor is dependent on the conditions of the environment and the receptor's exposure patterns, which are different site by site. Therefore site-specific information on environmental conditions, exposure pathways and rates, and

* Corresponding author. Tel.: +886-2-2363-0406; fax: +886-2-2392-8830.

E-mail address: hwma@ccms.ntu.edu.tw (H. Ma).

biological data is often needed to conduct site-specific risk assessment. A site-specific risk assessment may produce more accurate assessment than a generic assessment but increases complexity. Sampling, survey, investigation, and modeling are often involved in collecting site-specific data, and thus the amount of effort and time spent in the process may be too much to be ignored. The situation is additionally confounded by the stochastic nature of the natural world and by uncertainty due to incomplete information. Therefore it is desirable to identify important information so that the effort in data collection could be spent efficiently.

Sensitivity analysis methods based on partial derivatives of model output with respect to each input are often used to locate important information that significantly influences model prediction. A drawback of considering only simple sensitivity that measures the rate of change of model output with respect to variation in an input is that the degree of uncertainty is ignored. An input that has a small sensitivity but a large uncertainty may be as important as an input with a larger sensitivity but smaller uncertainty (Iman and Helton, 1988; Morgan and Henrion, 1990). Therefore uncertainty analysis methods that compare the importance of input uncertainties in terms of their contributions to uncertainty associated with the outputs are called forth. These methods include nominal range sensitivity method, standardized regression coefficient, partial regression coefficient, Pearson's correlation coefficient, rank correlation coefficient, and others (Iman and Helton, 1988; Morgan and Henrion, 1990; McKone, 1996; Mills et al., 1999). However, to facilitate solving a decision problem, such as the one that will be described later, the importance of information should be assessed related to the decision context. That means the knowledge about what kind of information is important in terms of its influence on decision-making is more valuable than the identification of important information based on its influence on model prediction.

This paper presents a case study of groundwater contaminated by chlorohydrocarbons. For this case problem, a baseline site-specific health risk assessment is performed, and a methodology is developed and implemented to assess the relative importance of various site-specific parameters in making risk management decisions. The case problem is described in the following section.

2. The groundwater contamination problem

The case study of this paper is about a site, located in northern Taiwan, where it was discovered that the soil and groundwater is contaminated by chlorinated hydrocarbons, primarily trichloroethylene (TCE) and tet-

rachloroethylene (PCE). These chemicals were used as degreasing solvents by a local factory but improper handling and disposal of wasted solvents led to leakage and subsequent contamination of soil and groundwater.

Chlorinated hydrocarbons are often detected in soil and groundwater because of their heavy use (for example, their use is estimated around 150,000 t annually in Taiwan) and unsound management of the raw or used substances. Public awareness of the issue is increasing as a result of disclosure of more contaminated sites. In the US, it was estimated in 1994 over a trillion dollars will be spent in the next 30 years on cleanup of contaminated soil and groundwater (US EPA, 1999). Among the 1430 national priorities list sites identified by the US EPA, TCE and PCE have been found in at least 852 and 771 sites, respectively. TCE and PCE may cause nervous system effects, and liver, lung, and kidney damage. And because PCE has been shown to cause liver tumors in mice and kidney tumors in male rats, it may reasonably be anticipated to be a carcinogen (ATSDR, 1996).

Extensive resources and efforts have been spent in cleaning up the site. The soil cleanup has been considered complete, but little success was reached for the groundwater, due to a dense-non-aqueous-phase-liquid's distinct physical characteristics that increase the difficulty of its identification and removal. At this point, the government is faced with a decision situation of determining whether the current groundwater contamination needs further remediation. Within the decision situation a vital problem is to identify information that is essential in reaching such a decision.

Based on the land use patterns and contaminant concentrations, the study site is divided into four divisions, each of which is assessed individually. The paper presents the study performed for one of the four divisions: the 80,000 m² of farming land with scattered residences and a large fishpond. According to the observations of 15 monitoring wells in this area, the mean and standard error of the contaminant concentrations are 0.12 and 0.02 mg/l for TCE, and 0.0005 and 0.0001 mg/l for PCE (Geomatrix et al., 1998). A site-specific risk assessment is performed and then important parameters are identified in terms of their influence on the decision regarding whether risk management measures should be initiated.

Section 3 presents the methodology for setting information priorities for the decision problem. Results and discussions are provided in Section 4, followed by a conclusion in Section 5.

3. Methodology

Site-specific multimedia risk assessment, uncertainty analysis, and decision criteria are combined to develop a

methodology for setting information priorities for the decision problem. In the literature related to groundwater risk assessment, researches that incorporate multimedia assessment are rare. Most researches consider the exposure due to use of groundwater only (McKone and Bogen, 1992; Andricevic and Cvetkovic, 1996; Maxwell et al., 1998; Lahkim and Garcia, 1999). The multimedia risk assessment method used in the case study is described below, and the procedures of setting information priorities are presented later.

3.1. Multimedia risk assessment

Site-specific multimedia risk assessment uses site-specific environmental and exposure information to establish the quantitative relationship linking the source of risk agents (in this case, the contaminants in the groundwater), transport and transfer through various environmental media, exposure to defined populations, and risk. In the case study, individual lifetime risk of developing cancer is used as the measure of health impact. The cancer risk is calculated by the following equation (US EPA, 1989):

$$\text{CancerRisk} = \sum_k [1 - \exp(\text{ADI}_k \times \text{CSF}_k)] \approx \sum_k \text{ADI}_k \times \text{CSF}_k, \quad (1)$$

where ADI_k is the average daily intake (mg/kg-d) of the contaminants by receptors through exposure route k (inhalation, ingestion, or dermal absorption), CSF_k is the cancer slope factor (kg-d/mg) of exposure route k , which is derived from dose-response studies.

In practice, the linearized portion is used for small values of $\text{ADI} \times \text{CSF}$. ADI is the amount of contaminants received by human bodies via all exposure pathways. It is the output of exposure assessment and is calculated based on the following equation (California EPA, 1993):

$$\text{ADI}_{ijk} = C_i \times \left[\frac{C_j}{C_i} \right] \times \left[\frac{\text{CR}_{jk}}{\text{BW}} \right] \times \frac{\text{EF} \times \text{ED}}{\text{AT}}, \quad (2)$$

where ADI_{ijk} (mg/kg-d) is the average daily intake from environmental medium i (air, soil or groundwater), exposure medium j (drinking water, food, etc.) and exposure route k (inhalation, ingestion, or dermal absorption), C_i is the pollutant concentration in environmental medium i , C_j is the pollutant concentration in exposure medium j , CR_{jk} is the contact rate of exposure medium j through exposure route k , EF is the exposure frequency, ED is the exposure duration, and AT is the averaging time. BW is the body weights of receptors.

Exposure assessment translates contaminant source terms into estimates of the amount of contaminant that comes into contact with an identified population. It consists of two steps: multimedia transfer and transformation modeling, and multiple-pathway exposure modeling. In the above equation, C_i , the contaminant concentrations in various environmental media, is estimated by the multimedia transfer and transformation modeling step. C_j/C_i represents the transfer between an environmental medium and an exposure medium, which is the purpose of the multiple-pathway exposure modeling step. Fig. 1 shows the transfer processes and exposure pathways involved in the two modeling steps of exposure assessment.

The procedures of conducting risk assessment for the case study are summarized as follows:

1. Starting with the concentration of TCE and PCE in the groundwater, calculate the diffusion of the contaminants from groundwater to outdoor air through the porous medium of soil (Johnson and Ettinger, 1991; ASTM, 1995).
2. Calculate contaminant concentrations in the indoor air due to the process that the vapor phase of the contaminants in the saturated zone diffuses through the porous medium of soil and the cracks in the foundation, floors, and walls of a building and into the indoor air (Johnson and Ettinger, 1991).
3. Calculate contaminant concentrations in soil due to irrigation and air deposition (Streng and Chamberlain, 1995).
4. Calculate contaminant concentrations in plant food, animal food, and personal air. The transfer processes between environmental media and exposure media include:
 - (i) Plant contamination due to root uptake, air deposition, and air-to-soil vapor transfer (Travis and Arms, 1988; McKone and Ryan, 1989; Bacci et al., 1990, 1992; US EPA, 1998).

Fig. 1. The multimedia transfer processes and multiple exposure pathways associated with a contaminated-groundwater site.

- (ii) Animal contamination due to plant and soil ingestion (Streng and Chamberlain, 1995; US EPA, 1998).
 - (iii) Fish contamination due to bioaccumulation from contaminated water supply (US EPA, 1998).
 - (iv) Contaminant concentrations in personal air due to volatilization of bathing water (McKone, 1987).
5. Using Eq. (2) to calculate the average daily intake received by receptors through various exposure pathways (The calculation of dermal absorption refers to US EPA, 1992). The intakes through the same exposure route (inhalation, ingestion, or dermal absorption) are summed.
 6. Using Eq. (1) to calculate individual cancer risk. The total cancer risk is the sum of risks received through all exposure routes.

The parameters needed for modeling environmental transfer and transformation include the physicochemical properties of the two chemicals and the environmental parameters that simulate the conditions of air, soil, and groundwater. When multiple-pathway exposure modeling is performed, the food chain parameters, biological and exposure factors, and chemical-specific biotransfer/bioconcentration factors are incorporated.

3.2. Combination of uncertainty analysis with decision criteria in identifying important information

In the study, a Monte Carlo technique is used to propagate parameter variance to produce the uncertainty associated with the risk estimate. Each of the input parameters that are considered uncertain is treated as a random variable with known or estimated probability density function (pdf). For each of the variables, one value is selected at random with respect to the associated pdf. The individual cancer risk is then calculated by using the multimedia risk assessment method with any sample set of input values. The sampling and calculation are repeated many times to produce the pdf of the risk estimate.

When a Monte Carlo technique is used to propagate variance, standardized regression coefficient, partial regression coefficient, Pearson's correlation coefficient, and rank correlation coefficient can be used to rank the importance of input parameters' uncertainty to risk estimation (Iman and Helton, 1988; Morgan and Henrion, 1990; McKone, 1996; Mills et al., 1999). Among these approaches, rank correlation coefficients are indicators of the degree of monotonic relationship between the sample values of the model prediction and those of the uncertain inputs. Therefore, rank correlation coefficients often work better to rank parameter contributions to uncertainty than the other approaches that are based

on linear relationships (Morgan and Henrion, 1990; McKone, 1996).

Another method that has been used to investigate parameter sensitivity is the regional sensitivity analysis (RSA) method developed by Spear and Hornberger (1980). In the method, random values are drawn from input parameters' pdfs and used to calculate a set of model outcomes. Each of the outcomes is assigned a binary classification based on whether or not the outcome is in agreement with specified behavior of the system. The method uses the degree of separation between the distribution of the set of random variables classified as "pass" and the set classified as "fail" to indicate the importance of parameters. The RSA method has been extended by incorporating a tree-structured density estimation technique to examine the complex interactions that exist between parameters (Spear et al., 1994). The extended RSA method is a useful approach to identify important parameters and model structures but involves computer intensive algorithms for multi-variable data analysis.

In the study, a rank correlation coefficient between each input parameter and associated risk output is computed to measure the importance of parameter uncertainty. Considering n pairs of samples from the output and a single input, denoted as y_i , x_i , for $i = 1$ to n , the simple correlation coefficient is calculated as follows:

Correlation coefficient

$$= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \times \sum_{i=1}^n (y_i - \bar{y})^2}},$$

where $\bar{x}$ and $\bar{y}$ are expected values of x and y , respectively.

When the sample values of x s and y s are replaced by the rank of those values, the above equation calculates the rank correlation coefficient.

Although rank correlation coefficients provide a method of assessing parameter importance to model prediction, in order to identify information important to decision-making, decision criteria need to be incorporated into the uncertainty analysis practice. The procedures are as follows:

1. Perform the Monte Carlo simulation and calculate the rank correlation coefficients. Then the rank correlation coefficients are squared and normalized to calculate the relative percentage of the total variance attributable to each of the parameters.
2. Select a parameter that is identified to have major contribution to the overall variance. Fix it at its plausible minimum value and perform the Monte Carlo simulation to find the risk value at the 95% cumulative distribution level.

3. Fix the value of this parameter at its plausible maximum and perform the Monte Carlo simulation to find the risk value at the 95% cumulative distribution level.
4. Examine whether the resulting range of risk generated above extends across the target level specified in the decision criteria. If it is the case, the variation of the value of the parameter may affect the result of decision; otherwise the value of the parameter would not influence the decision, even though the parameter may have major contribution to the overall variance of risk estimation.
5. Repeat steps 2–4 for all the parameters that are identified in step 1 as having major contribution to the overall variance.

4. Results and discussion

Since 95% cumulative probability distribution level is normally considered in risk-based decision-making, the results of risk values at the 95% cumulative probability distribution level are presented and discussed in the section.

The total individual cancer risk at the 95% cumulative distribution level is $1.96\text{E}-5$. The contribution of each parameter's uncertainty to the variance of the risk estimate based on rank correlation coefficients is shown in Table 1. Seven parameters with contribution greater than 1% are listed and are considered the important parameters for the multimedia risk estimation. Those parameters should be treated more carefully, probably by conducting further research to obtain more information about the parameters in order to obtain more accurate risk estimation.

Improving model prediction is often not the purpose of identifying important information in a decision problem. Instead, the collection of information is aimed to facilitate the making of the decision. Therefore, usually whether variation in an input variable would make a difference to decision-making is of primary interest. The

methodology described previously combines uncertainty analysis and decision criteria to obtain further insights on the importance of the parameters for the decision problem of determining whether the contamination should be subject to risk management. Table 2 shows the results of following steps 1–3 stated in the methodology section for the seven parameters identified influential to model prediction. In the case study, it is decided that the screening decision is based on the health risk only. Furthermore, $1\text{E}-5$ at 95% cumulative probability distribution level is used as the target level below which no risk management effort is required. It is found that only four of the seven parameters have the target risk level, $1\text{E}-5$, falling between the two extremes. These four parameters are TCE concentration, vegetable yield, deposition interception fraction of vegetables, and plant surface loss constant. In particular, the two parameters, ingestion rate of drinking water (with 20% contribution to the overall variance) and irrigation rate (with 12% contribution to the overall variance), are not shown to affect the decision. This indicates that the parameter with major contribution to the overall variance is not necessarily the parameter that influences the decision and should be treated with priority.

Although the health-based criterion is used in the case study, in general, the decision criteria may include utility-based (e.g., considering both risk and cost of implementing risk management measures such as remediation), technology-based (applying the best available technology), and right-based (health risk is the only consideration) criteria (Morgan and Henrion, 1990; Thompson and Graham, 1996). The selection of decision criteria will depend on the situation (including political and socioeconomic factors), although in screening whether actions should be initiated, the last criterion, the right-based criterion, is normally used.

Even when the health risk is used as the determinant of a screening decision, there may be different policies

Table 1
Contribution of each parameter to the variance of the total risk estimate

Parameter	Variance contribution (%)
TCE concentration	24
Ingestion rate of drinking water	20
Yield of vegetables	12
Irrigation rate	11
Fraction of deposition retained on plant surface	6
Plant surface loss constant	3
Bathroom size	2

Table 2
95% Cumulative distribution risk value with a parameter set at its plausible lower bound or upper bound

Parameter	Total risk	
	Lower bound	Upper bound
TCE concentration	$2.72\text{E}-6$	$2.82\text{E}-5$
Ingestion rate of drinking water	$1.13\text{E}-5$	$3.21\text{E}-5$
Yield of vegetables	$9.60\text{E}-6$	$6.47\text{E}-5$
Irrigation rate	$1.15\text{E}-5$	$5.59\text{E}-5$
Deposition interception fraction of vegetables	$9.20\text{E}-6$	$3.59\text{E}-5$
Plant surface loss constant	$9.50\text{E}-6$	$2.34\text{E}-5$
Bathroom size	$1.53\text{E}-5$	$2.01\text{E}-5$

for acceptable risk than specifying the acceptable level at a particular percentile of the risk distribution (95% in this case). For example, the acceptable risk may be defined as a distribution itself, inequality constraints on specified moments of the risk distribution, or inequality constraints on several percentiles of the risk distribution, as detailed in Thompson and Graham (1996). No matter which policy is used, the above results indicate that the information essential to decision making can be located by combining Monte Carlo simulation, rank correlation coefficients, and specified decision criteria. This method provides better basis for the decision than only ranking parameters based on uncertainty contribution.

5. Conclusion

In the case study of the site with contaminated groundwater, site-specific risk assessment taking into account multimedia transfer and multiple exposure pathways is considered an important basis of decision concerning whether actions should be in place. The site-specific multimedia risk assessment may produce more accurate risk estimate than a generic risk assessment practice but requires more information that is often complicated by the associated uncertainties. Therefore the knowledge regarding the most important information that would affect the decision will be beneficial in making data collection planning. To that end, the sensitivity analysis and uncertainty analysis methods that consider parameters' effects to model prediction need to be extended to incorporate decision criteria to identify information essential to the decision. The results demonstrate that a parameter with major contribution to model uncertainty and much influence on model prediction may not affect the making of a decision. The integration of multimedia risk assessment, Monte Carlo simulation, and decision criteria is a useful and simple method of setting information priorities.

Acknowledgements

We would like to thank the National Science Council of Taiwan for the financial support of this work under contract NSC89-2621-Z-002-020.

References

- American Society for Testing and Materials, 1995. Standard guide for risk-based corrective action applied at petroleum sites. E1739-95. West Conshohocken, PA.
- Agency for Toxic Substances and Disease Registry (ATSDR), 1996. Toxicological profile for tetrachloroethylene (Update). Atlanta, GA.
- Andricevic, R., Cvetkovic, V., 1996. Evaluation of risk from contaminants migrating by groundwater. *Water Resources Research* 32, 611–621.
- Bacci, E., Calamari, D., Gaggi, C., 1990. Bioconcentration of organic chemical vapors in plant leaves: experimental measurements and correction. *Environmental Science and Technology* 24, 885–889.
- Bacci, E., Cerejeira, M.J., Gaggi, C., 1992. Chlorinated dioxins: volatilization from soils and bioconcentration in plant leaves. *Bulletin of Environmental Contamination and Toxicology* 48, 401–408.
- Barken, C., Kroll, S., Sterret, R., Lynch, J., Waggner, J., Kostecki, P., 1996. State by state summary of RBCA approaches. *Soil Groundwater Clean-up*, April, pp. 41–53.
- California EPA, 1993. CalTOX: A multimedia total exposure model for hazardous-waste sites, Sacramento, CA.
- Geomatrix Consultants, Inc., Beak Consultants, Inc., Kueper & Associates, 1998. Groundwater Remediation Technical Practicability Evaluation Report, Taoyuan Site, Taiwan.
- Iman, R.L., Helton, J.C., 1988. An investigation of uncertainty and sensitivity analysis techniques for computer models. *Risk Analysis* 18, 71–90.
- Johnson, P.C., Ettinger, R.A., 1991. Heuristic model for predicting the intrusion rate of contaminant vapors into buildings. *Environmental Science and Technology* 25, 1445–1452.
- Lahkim, M.B., Garcia, L.A., 1999. Stochastic modeling of exposure and risk in a contaminated heterogeneous aquifer 1: Monte Carlo uncertainty analysis. *Environmental Engineering Science* 16, 315–328.
- Maxwell, R.M., Pelmulder, S.D., Tompson, A.F.B., Kastenber, W.E., 1998. On the development of a new methodology for groundwater-driven health risk assessment. *Water Resources Research* 34, 833–847.
- McKone, T.E., 1987. Human exposure to volatile organic compounds in household tap water: the indoor inhalation pathway. *Environmental Science and Technology* 21, 1194–1201.
- McKone, T.E., 1996. Alternative modeling approaches for contaminant fate in soils: uncertainty, variability, and reliability. *Reliability Engineering and System Safety* 54, 165–181.
- McKone, T.E., Bogen, K.T., 1992. Uncertainties in health-risk assessment: an integrated case study based on tetrachloroethylene in California groundwater. *Regulatory Toxicology and Pharmacology* 15, 86–103.
- McKone, T.E., Ryan, P.B., 1989. Human exposure to chemicals through food chain: an uncertainty analysis. *Environmental Science and Technology* 23, 1154–1163.
- Mills, W.B., Lew, C.S., Hung, C.Y., 1999. Sensitivity analysis of concentration and risk predictions in the PRESTO and MMSOIL multimedia models: regression technique assessment. *Risk Analysis* 19, 511–525.
- Morgan, M.G., Henrion, M., 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, New York.
- Spear, R.C., Hornberger, G.M., 1980. Eutrophication in peel inlet, II, identification of critical uncertainties via generalized sensitivity analysis. *Water Research* 14, 43–49.
- Spear, R.C., Grieb, T.M., Shang, N., 1994. Parameter uncertainty and interaction in complex environmental models. *Water Resources Research* 30, 3159–3169.

- Streng, D.L., Chamberlain, P.J., 1995. The multimedia environmental pollutant assessment system (MEPAS): exposure pathway and human health impact assessment models. Prepared for US DOE at Battelle.
- Thompson, K.M., Graham, J.D., 1996. Going beyond the single number: using probabilistic risk assessment to improve risk management. *Human and Ecological Risk Assessment* 2, 1008–1034.
- Travis, C.C., Arms, A.D., 1988. Bioconcentration of organics in beef, milk, and vegetation. *Environmental Science and Technology* 22, 271–274.
- US EPA, 1989. Risk assessment guidance for superfund. Human Health Evaluation Manual, vol. 1, (Part A), Office of Emergency and Remedial Response.
- US EPA, 1992. Dermal exposure assessment: principles and applications. Office of Research and Development, Washington, DC.
- US EPA, 1998. Methodology for assessing health risks associated with multiple pathways of exposure to combustor emissions. Cincinnati, OH.
- US EPA, 1999. Safe drinking water Act, Section 1429: Groundwater Water Report to Congress, Washington, DC.