

available at www.sciencedirect.comwww.elsevier.com/locate/scitotenv

Assessment of the value of reducing uncertainty by sampling in a groundwater remediation system

Hwong-wen Ma*, Chih-Chen Chang

Graduate Institute of Environmental Engineering, National Taiwan University, 71 Chou-Shan Rd., Taipei, Taiwan 106

ARTICLE INFO

Article history:

Received 14 February 2007

Received in revised form

14 April 2008

Accepted 19 April 2008

Keywords:

Value of information

Random field generation

Conditional simulation

Multiple-realization

management model

Groundwater remediation

ABSTRACT

The success of cleaning up a contaminated site depends on the degree of knowledge of the site's characteristics. When there is much uncertainty associated with the knowledge, the uncertainty regarding whether the remediation will work increases consequently. It is therefore essential to know how much reduction of uncertainty is needed for the purpose of designing a successful and reliable remediation system. The understanding of the site characteristics is basically increased by site investigation, and thus the uncertainty is decreased by sampling information. This study develops a method to evaluate the value of reducing uncertainty by sampling the hydrogeological parameters in a groundwater remediation system. Hydraulic conductivity being taken as an example of the site characteristics, random field generation and conditional simulation are coupled to obtain a range of hydraulic conductivity fields based on the sampling outputs. A multiple-realization management model that incorporates a chance constraint of health risk is used to find the lowest remediation cost under specific remediation criteria of risk through genetic algorithm. The remediation cost, which is expected to decrease with collection of more samples, serves as the measure of the value of uncertainty reduction by sampling. A case study shows that the variation of the hydraulic conductivity fields among the potential sites as well as the remediation cost is reduced as a result of increase of samples. It also shows that the risk after remediation decreases with the collection of more samples, which implies that the reduction of risk can also be used to assess the value of sampling.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Because of inappropriate disposal of hazardous wastes or materials for a long time, the problems of groundwater contamination have emerged and raised serious concern. Some of the contaminated sites may pose threat to the environment and human health, in addition to causing deterioration of environmental quality and loss of utility of the environmental resources. When the potential or realized impact is considered unacceptable, remediation measures are often required to remove the contaminants. To clean up a site with contaminated groundwater has to consider a wide range of factors, such as the hydrogeology and pollutant properties.

Those factors affect the behavior of the pollutants in the environment, leading to distinct spatial and temporal distribution of the pollutants around the sites and hence the magnitude and pattern of resultant impact. The knowledge of those factors will therefore influence the design of remediation plans and the results of the remediation work. A remediation scheme is usually established for a contaminated site with characteristics that are specified but not precisely known. When there is insufficient understanding and thus much uncertainty associated with the characterization of the relevant factors, the remediation may be either ineffective, resulting in unsatisfactory reduction of risk, or inefficient, incurring unnecessary increase of cost. Whether a remedia-

* Corresponding author. Tel.: +886 2 23630406; fax: +886 2 23928830.

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

tion system works depends greatly on the degree of understanding of the characteristics of the contaminated site. It is therefore essential to know how precise the knowledge about the characteristics is sufficient and how much reduction of uncertainty associated with the estimation of the characteristics is needed for the purpose of designing a successful and reliable remediation system. The understanding is basically increased by site investigation and thus the uncertainty can be decreased by sampling information.

Part of the studies in the literature concerning the sampling for the groundwater system has been focused on pollutant monitoring. Geostatistical techniques have been considerably involved in designing sampling networks. The magnitude of the estimated values and reduction of Kriging variance are usually used as a basis of optimal sampling strategies (Rouhani and Hall, 1988; Van Groenigen, 2000). Groundwater transport modeling has also been combined with Monte Carlo simulations to generate thousands of realizations of pollutant distribution. Based on the realizations, an optimal monitoring network is designed to maximize the probability of detecting the contaminant plume and minimize the expected volume of contaminated groundwater at the time of detection (Meyer and Brill, 1988; Meyer et al., 1994; Storck et al., 1997). Cost minimization and the estimation error of contamination mass interpolation have also been considered in locating monitoring wells by combining a fate-and-transport model and genetic algorithm (Reed et al., 2000).

In the literature of dealing with the sampling design of aquifer parameters, uncertainty analyses of parameter estimate and model prediction of contaminant concentrations are combined with optimization to find the sampling strategy that leads to the greatest reduction in model prediction uncertainty subject to budget constraint (Wagner, 1995). When aquifer parameters' uncertainties are considered under the framework of groundwater management such as remediation, two major approaches have been used (Aly and Peralta, 1999). One is that parameter uncertainty and the resultant contaminant concentration uncertainty are built into a management model. A chance-constrained management model, which ensures water quality standards are met at a specified reliability level, is used to relate reduction in model prediction uncertainty to the reduction of groundwater management costs. The model implies that at a specified reliability level, the greater the prediction uncertainty, the lower the expected value of the concentration resulting from remediation is required to be (Wagner and Gorelick, 1987; Wagner, 1999). The shortcoming is that the spatial structure of the regional variables is not considered and that the relationship between sampling and the contaminant concentrations at the checkpoints has to be identified based on some amount of acquired data in advance.

The other way of incorporating the uncertainty in the aquifer's properties into groundwater management is the use of a Monte Carlo approach. Multiple realizations of an aquifer parameter's field are generated randomly based on the spatial structure derived from sampling information. The multiple fields are examined simultaneously, aiming at identifying an optimal management strategy that satisfies concentration criteria for each of the fields (Wagner and Gorelick, 1989; James and Gorelick, 1994). In this approach, the Kriging method and a random field generation method such as the

Turning Bands method (Mantoglou and Wilson, 1982) and Hydro_Gen (Bellin and Rubin, 1996) are combined to perform conditional simulation to generate a series of equally probable random fields. Those conditionally-simulated random fields preserve the measured values at sampled locations and reproduce accurately the prescribed spatial statistics (Delhomme, 1979; Warner et al., 2006; Bau and Mayer, 2007; Mantoglou and Kourakos, 2007; Ricciardi et al., 2007).

The purpose of this study is to develop a method to evaluate the value of reducing uncertainty by sampling aquifer parameters in a groundwater remediation system. Such evaluation would provide information of the worth of sampling aquifer parameters for the purpose of increasing the efficiency and effectiveness of remediation, and serve as the basis for guiding site characterization. This study adopts the concept of multiple realizations of an aquifer parameter's field, employing the variability among the realizations as the manifestation of the uncertainty of the parameter's field of a contaminated site. The idea of chance constraint is also integrated to form a chance-constrained multiple-realization management model. Through the management model, the effect of such uncertainty on remediation is estimated and the value of sampling information evaluated. The objective of the management model is to find the lowest-cost remediation strategy that satisfies the risk acceptability criteria. In this study, it is the health risk caused by the contaminated site, as opposed to the contaminant concentration traditionally, that serves as the remediation criteria in the chance constraint of the management model, because the goal of remediation is to reduce the risk in an economical way. The rationale is that as the knowledge of the spatial distribution of the parameter is enhanced by sampling, the variability among the field realizations will be decreased, which in turn lead to the reduction of cost of meeting the risk criteria at a specified reliability level. The amount of cost reduction can therefore be used to measure the value of the uncertainty diminished by sampling.

A site contaminated by perchloroethylene (PCE) is used as a case to test the evaluation methodology. The hydraulic conductivity is used as an example of the spatial hydrogeologic attributes due to its significance to contaminant transport and high cost of measurement (Ricciardi et al., 2007). The random field generator Hydro-Gen is combined with conditional simulation to produce numerous realizations of the hydraulic conductivity field. MODFLOW and MT3DMS are used to model the transport of PCE as a result of a remediation scheme. The next section describes the methodology to assess the value of sampling information in a groundwater remediation system in terms of risk and cost. The application of the methodology to a case study is discussed in Section 3. Section 4 presents results and discussion of the case study, followed by the conclusion in Section 5.

2. Materials and methods

The methodology of estimating the value of sampling information is mainly composed of two stages. In the first stage, sampling is conducted on the contaminated site. Based on the revealed sampling information, a random field generation method (Hydro_Gen is used in this study) combined with

conditional simulation is performed to generate several random fields (say, there are N simulated sites) that represent the range of possible site characteristics. As a spatially distributed random field generator for correlated properties, Hydro_Gen has the ability to reproduce accurately the prescribed spatial statistics (Bellin and Rubin, 1996). Compared to Kriging technique that aims to estimate the values at particular locations based on measurements, random field generation is used to produce the expected values and variances at the locations and simulate the potential random fields, which have the same spatial structure of the regionalized variables and reflect the spatial variations and uncertainty of estimation. To consider the effect of sampling on the knowledge of site characteristics, conditional simulation is further conducted to ensure that each sampling position has the same actual sampling value in all realizations.

Let $z(x)$ and $z^*(x)$ be the true value at point x and Kriging estimate based on actual sample values, respectively. Before conditioning, the true value $z(x)$ is not available if x is not a sample point and the Kriging error $z(x) - z^*(x)$ remains unknown. The true value can be decomposed as: $z(x) = z^*(x) + [z(x) - z^*(x)]$. Kriging can be performed similarly using the values taken by a given field realization at the sample locations as data. Let $s(x)$ be a value at a field realization and $s^*(x)$ is a Kriged estimated value derived from Hydro_Gen simulated values corresponding to sampling locations. The known error term $s(x) - s^*(x)$ is substituted for $z(x) - z^*(x)$, and then the conditional simulation value is equal to $z^*(x) + s(x) - s^*(x)$. At the point of sampling, $s(x)$ is equal to $s^*(x)$ and thus its final simulated value is the actual sample value. For further discussion refer to Delhomme (1979).

In the second stage, groundwater transport modeling is combined with an optimization algorithm to locate the optimal remediation strategy with the lowest cost and satisfied criteria concerning risk across the multiple sites. In this study, the genetic algorithm (Mitchell, 1996) is used to generate a candidate remediation strategy that is applied to the N simulated sites. The U.S. Geological Survey modular finite-difference groundwater flow model (MODFLOW) (Harbaugh et al., 2000) and the Modular 3-Dimensional Multi-Species Transport model (MT3DMS) (Zheng and Wang, 1999), are then performed on the N situations and produced N outputs of contaminant concentration distribution. The N health risks as a result of the N outputs of concentration fields are subsequently examined by genetic algorithm against the defined remediation criteria. Various remediation strategies are generated and examined in such manner until the lowest-cost strategy is located.

In the traditional multiple-realization management model, the remediation is deemed successful only when all the N sites satisfy the specified criteria concerning water quality standards. In order to take into account the variation among the sites, which is reflecting the degree of understanding of the site characteristics, the chance-constrained model is adapted as follows:

$$\begin{aligned} & \text{Min Cost} \\ & \text{Subject to} \\ & \text{Avg}(\text{Risk}_i) + F^{-1}(\pi) \times \text{Sd}(\text{Risk}_i) \leq \text{Risk}^* \end{aligned}$$

where $\text{Avg}(\text{Risk}_i)$: the average of the N random fields' risk ($i = 1 \sim N$)

$F^{-1}(\pi)$: value of the standard normal cumulative distribution corresponding to reliability level π

$\text{Sd}(\text{Risk}_i)$: the standard deviation of the N random fields' risk
 Risk^* : the target risk level.

At a specified reliability level and target acceptable risk, when the variation among the realizations is smaller as a result of sample increase, the expected risk achieved by remediation can be allowed to be larger and therefore the remediation cost can be reduced. The two-stage procedure is repeated for various numbers of sampling points, and the change of the remediation cost with different numbers of samples can then serve as a measure of the value of the sampling information.

Looking more deeply from microscopic perspective, the value of some parameter at a point, say point A, falls within the range of $v_1 \sim v_2$ among N possible random fields before sampling and conditional simulation are conducted. When conditional simulation is performed based on the value collected at a sampling position, the range of the parameter's value at point A becomes $v_1^* \sim v_2^*$. If point A happens to be the sampling point, $v_1^* = v_2^* = \text{sampling value}$; if point A is somewhere near the sampling point, the width of range $v_1^* \sim v_2^*$ will be narrower than that of range $v_1 \sim v_2$. The narrowing effect fades with the distance between point A and the sampling position and magnifies as more points are sampled. That the range of variation becomes smaller means that the values at the same position among the N random fields become increasingly similar. Therefore when the N sites are to be managed with success, the incurred cost will decrease with the decrease of difference among those sites. By this concept, the reduction of uncertainty could be represented by the reduction of remediation cost to find the value of sampling information.

3. Case study

A site located at the southwest of Taoyuan city in northern Taiwan and contaminated primarily by perchloroethylene (PCE) is used as a case to demonstrate the evaluation methodology. PCE and other solvents were used as degreasing agents by an electric equipment production factory and deposited into the first aquifer over a period of ten years before discovery of the issue. Because of a long period of improper management, the discovery of heavy PCE contamination forced the government to face the problem of the long-term health risks of local residents.

This site has attracted national attention because serious soil and groundwater contamination has been detected and excessive death and cancer cases among the employees have been reported. The lawsuit filed by the employees against the company that improperly disposed of the chlorinated solvents is still on-going. A large amount of resources has been spent to clean up the site. Although soil contamination has been removed, groundwater remediation is not successful. Insufficient grasp of the site characteristics is a major cause of the failure. This study simulates the contamination history that includes a pump-and-treat remediation, and hydraulic conductivity (K) is used as an example of concerned site characteristics. The following describes the steps of testing the methodology through the case study and the associated settings for the case study.

3.1. The setting of a true hydraulic conductivity field

Based on the investigation report of the case site (Geomatrix et al., 1998) and the assumption of the spatial correlation structure (Hoeksema and Kitanidis, 1985; Aly and Peralta, 1999; Reed et al., 2000), the expected value and variance of $\ln(K)$ are set to be 2.385 m/day and 0.223, respectively. Hydro-Gen is then used to generate a field of hydraulic conductivity that is assumed to be the true field.

3.2. Modeling of contaminant transport in the groundwater

According to the site investigation (Geomatrix et al., 1998), we focus on a 120,000 m² (400 m*300 m) region including a farming land with scattered residences. Fig. 1 is a planar view of the simulated aquifer. In the modeling process, the aquifer is confined with constant head boundaries specified at the bottom (up-gradient) and top (down-gradient). The left and right boundaries of the aquifer are impermeable. Two hot spots are located at (row 36, column 13) (denoted as (36, 13)) and (28, 24) while the origin of the coordinate is located at the northwest corner. The PCE concentration at the hot spots has been determined to be 5.9 ppm and 1.6 ppm, respectively. To simulate the pollution history, 10 years of pollutant trans-

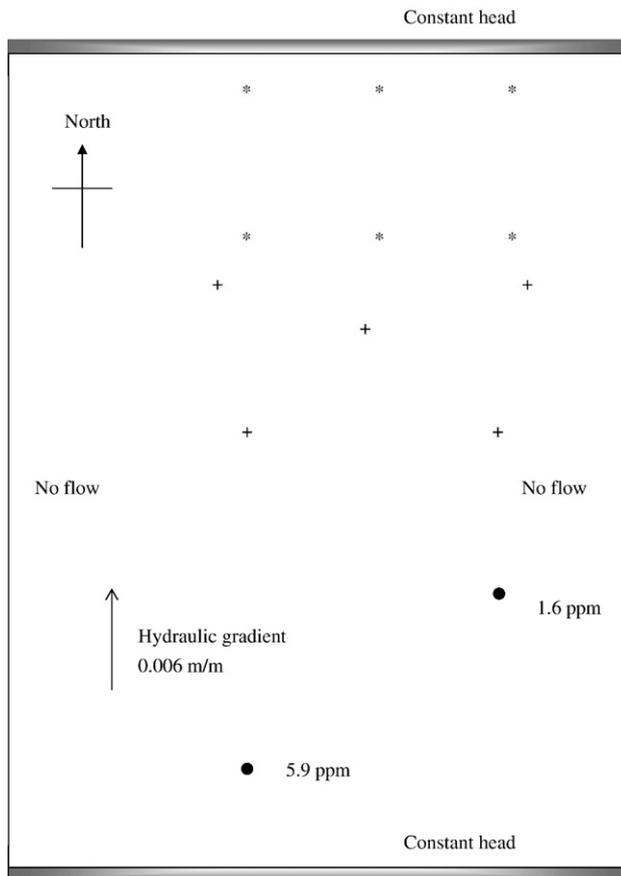


Fig. 1 – Position of pump wells, check points and contaminant source (+: pump wells; *: check points; •: source of contamination).

Table 1 – Hydrogeologic parameters used in groundwater transport modeling

Parameter	Value
Effective porosity	0.25
Hydraulic gradient	0.006 m/m (North)
Longitudinal dispersivity	3.00 m
Transverse dispersivity	0.90 m
Bulk density	1.7 g/cm ³
Specific yield	0.2
Fraction of organic carbon	0.01

port is simulated by MODFLOW and MT3DMS. The aquifer properties used in MODFLOW and MT3DMS are shown in Table 1. The simulated output is set as the initial pollutant concentration field before remediation takes place.

3.3. Generation of multiple random fields based on sampling

A sample is taken randomly and then Hydro_Gen is combined with conditional simulation to obtain N realizations of the hydraulic conductivity field; 5 programs of sampling are tested and N is set to be 30. Because in reality the true mean and variance of $\ln(K)$ are not known but could be obtained by consulting historical investigation information, they are assumed to be 2.3 and 0.25, respectively, in the test case. The range parameter used in Hydro_Gen is set to be 20 m.

3.4. Modeling of groundwater transport as a result of remediation work

MODFLOW and MT3DMS are used to calculate the PCE concentration field after 6 months of pump-and-treat (the actual period of pump-and-treat period was from March to June of 1997). Using initial contaminant concentrations as inputs, the groundwater transport modeling is conducted at each of the 30 hydraulic conductivity field realizations and for various pumping rates. Given a number of samples, 30 PCE concentration fields are obtained for a particular pumping rate. 5 pumping wells are located at (20, 13), (20, 23), (14, 17), (12, 11), and (12, 25); the 6 checkpoints, (1, 13), (1, 18), (1, 24), (10, 13), (10, 18), and (10, 24), produce well water for subsequent use (as shown in Fig. 1). The resultant contaminant concentrations estimated at the checkpoints are subsequently used in risk assessment.

3.5. Identification of the optimal remediation strategy

Genetic algorithm is used to identify the optimal remediation strategy (Palisade Corporation, 2001). As an example, the pumping rate is the only decision variable to be decided in the case study. The pumping rate could not exceed 1600 m³/day to avoid negative head. The decision goal is to minimize the cost of pump-and-treat while satisfying the constraint of risk. The residents had used the groundwater for daily activities, including drinking and irrigation, before the contamination was discovered. The health risk is estimated by combining the

contaminant concentrations in the checkpoints with the exposure pathways. To simplify the example, only the exposure of groundwater drinking is considered and the resulting carcinogenic risk represents the primarily concerned health impact (Ma, 2002). The optimization framework is formulated as follows:

$$\text{Min } C_p + C_t$$

$$\text{Subject to } \text{Avg}(\text{Risk}_i) + F^{-1}(99\%) \times \text{Sd}(\text{Risk}_i) \leq 1E - 6, \quad i = 1-30$$

$$C_p = \sum_j \left[14,000 + 39,355 \times (Q_j/86,400)^{0.453} + 1.5 \times Q_j \right]$$

$$C_t = \sum_j \left[39,000 \times (Q_j/1635) + 16.64 \times Q_j \right]$$

$$\text{Risk}_i = \text{SF} \times \frac{C \times \text{CR} \times \text{EF} \times \text{ED}}{\text{BW} \times \text{AT}}, \quad i = 1-30$$

$$0 \leq Q_j \leq 1600 \text{ m}^3/\text{day}, \quad j = 1-5$$

where Risk_i : the risk of drinking groundwater from the six checkpoints ($i=1-30$)

$\text{Avg}(\text{Risk}_i)$: the average of the 30 fields' risk ($i=1-30$)

$F^{-1}(99\%)$: value of the standard normal cumulative distribution corresponding to reliability level 99%

$\text{Sd}(\text{Risk}_i)$: the standard deviation of the 30 fields' risk

C_p : cost of pumping in US\$ (McKinney and Lin, 1996)

C_t : cost of treatment using activated carbon in US\$ (Nicholas, 1997)

Q_j : pumping rate of well j in unit of m^3/day , $j=1-5$

C : average concentration of the six checkpoints in unit of mg/L .

CR : consumption rate of drinking water in unit of L/day

EF : exposure frequency in unit of day/year

ED : exposure duration in unit of year

BW : body weight in unit of kg

AT : averaging time

SF : cancer slope factor in unit of kg-day/mg

Different remediation technologies can be assessed as long as the optimization framework contains the cost and risk as functions of the technologies. An alternative decision framework can also be used to evaluate the value of uncertainty reduction by sampling. The objective and constraint in the management model can be switched; i.e. the objective of the remediation is to minimize the risk associated with the contaminated site subject a budget constraint. The optimization framework can be formulated as follows.

$$\text{Min } \text{Avg}(\text{Risk}_i) + F^{-1}(99\%) \times \text{Sd}(\text{Risk}_i) \leq 1E - 6$$

$$\text{Subject to } C_p + C_t \leq B$$

$$C_p = \sum_j \left[14,000 + 39,355 \times (Q_j/86,400)^{0.453} + 1.5 \times Q_j \right]$$

$$C_t = \sum_j \left[39,000 \times (Q_j/1635) + 16.64 \times Q_j \right]$$

$$\text{Risk}_i = \text{SF} \times \frac{C \times \text{CR} \times \text{EF} \times \text{ED}}{\text{BW} \times \text{AT}}, \quad i = 1-30$$

$$0 \leq Q_j \leq 1600 \text{ m}^3/\text{day}, \quad j = 1-5$$

where B is the budget (set to be US\$300,000 in this example).

4. Results and discussions

In the case study for testing the methodology, 5 programs of sampling are evaluated. The effects of the sampling are presented and discussed as follows.

4.1. Variation of a regionalized variable with sampling numbers

The effect of sampling on the understanding of a particular point can be revealed by examining the point's characteristics in the 30 sites simulated based on a given number of sampling. Using point (50, 300) as an example, Fig. 2 shows that the variation among the realizations has a decreasing trend along with increase of samples. Table 2 provides the statistics of $\ln(K)$ at specific points of the 30 simulated sites under circumstances of different sampling numbers. Conditional simulation therefore shows that, from a microscopic perspective, the uncertainty of the estimations at individual points is reduced due to sampling.

4.2. Change of risk variation with sampling numbers

To see the macroscopic effect of conditional simulation on the variation of simulated random fields, the risks resulting from the 30 simulated random fields that went through the same 6 months without remediation are calculated. Fig. 3 shows the variation (represented by coefficients of variation) of the risks among the 30 simulated random fields corresponding to the number of collected samples used in the conditional simulation. The trend shows that increase of samples will reduce the variation of risks across the multiple simulated sites. The combination of the previous two figures implies that the different realizations become more similar as more samples are collected and subsequently result in more similar risk estimates.

In a sampling program, adding one sample does not necessarily decrease the variation. It is found that when a very large value is obtained from the additional sample, the

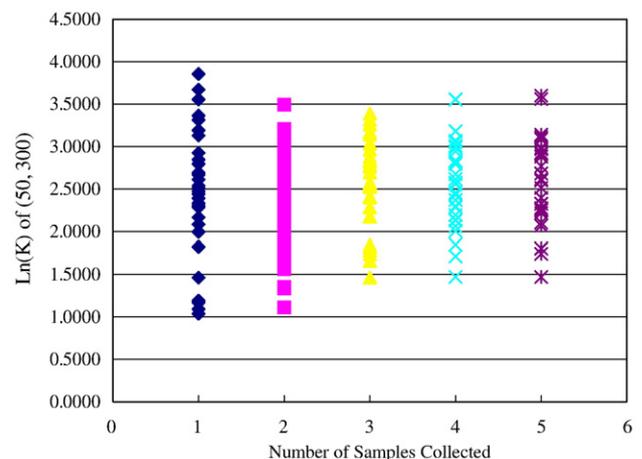


Fig. 2 – $\ln(K)$ at point (50,300) in the multiple simulated fields influenced by the number of samples collected.

Table 2 – Statistics of $\ln(K)$ at individual positions influenced by the number of samples

No. of samples	Statistics	Position (X,Y)				
		(50,300)	(50,100)	(150,200)	(250,300)	(250,100)
1	Mean	2.4644	2.3525	2.2835	2.3588	2.3387
	CV	0.3093	0.3266	0.2689	0.3345	0.2954
2	Mean	2.3762	2.2646	2.1747	2.2708	2.2533
	CV	0.2556	0.2680	0.1925	0.2866	0.2783
3	Mean	2.5149	2.3594	2.2516	2.3655	2.3453
	CV	0.2108	0.2308	0.1913	0.2649	0.2479
4	Mean	2.5958	2.6196	2.3347	2.4676	2.4445
	CV	0.2010	0.1423	0.1864	0.2387	0.2467
5	Mean	2.6025	2.6240	2.3416	2.4778	2.4527
	CV	0.1941	0.1418	0.1822	0.2322	0.2337

resultant risk variation would become greater. In Fig. 3, the third sampling $\ln(K)$ value in program 1 and the second sampling $\ln(K)$ values in program 3 are 2.9283 and 2.8598, respectively, which are much greater than the mean value, 2.3. The effect of the extreme sampling values would be diluted when more samples are collected.

4.3. Remediation cost and sampling numbers

Fig. 4 depicts the relationships between the remediation costs and number of samples collected. The remediation cost is calculated by the average cost of the five sampling programs. Fig. 4 demonstrates that the cost of remediation diminishes with the increase of sampling. The decrease of cost is expected to level off along with the increase of sampling numbers. This figure thus represents the value of sampling information by measuring the magnitude of the decrease of remediation cost.

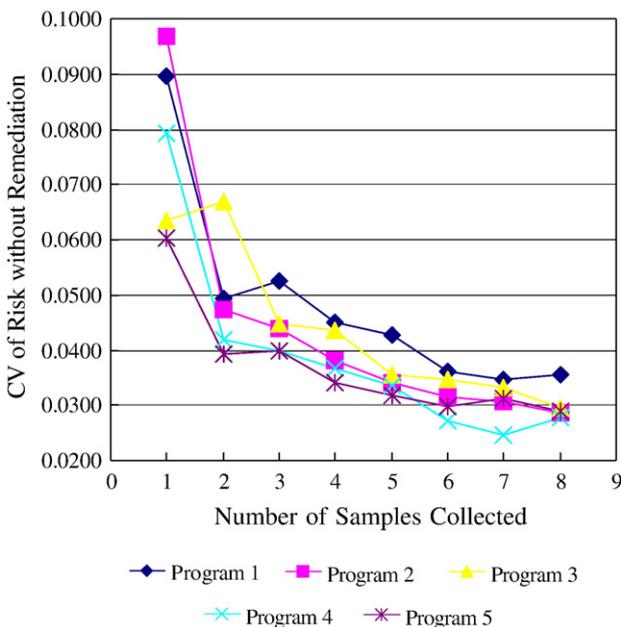


Fig. 3 – The effect of sampling numbers on the CV of risk without remediation.

4.4. The effects of the random field generator’s parameters

Besides sampling values, the mean and variance of the spatial structure and the range parameter would affect the outputs of random field generation.

4.1.1. The effect of mean and variance on the value of information

It is found that the random fields generated would be different with different values of the mean $\ln(K)$ used in Hydro_Gen. However, when the conditional simulation is performed based on sampling, the resulting fields become identical regardless of the mean value used. However, when different values of variance are used, the resultant curves of remediation cost with increase of samples differ. Fig. 5 shows the curves of average remediation cost with number of samples using two values of variance, 0.25 and 0.3. It indicates that the random field generation using higher values of variance would lead to higher remediation cost for a given number of samples and steeper trend of cost decrease with increase of samples. Therefore the value of sampling as represented by reduction of cost is larger and thus more samples acquisition tend to be needed; selecting a higher value of the variance is hence considered conservative.

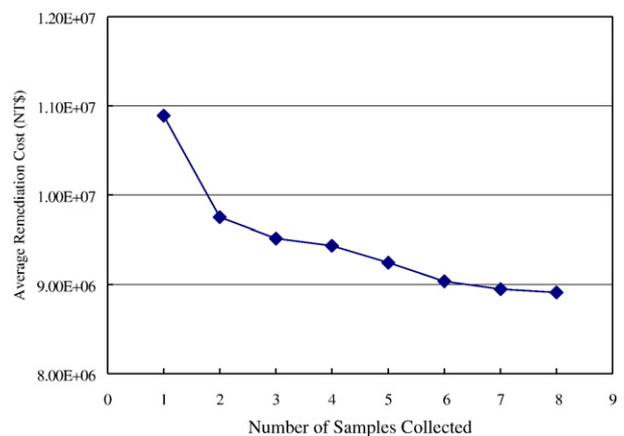


Fig. 4 – The relationships between the average remediation costs and number of samples collected.

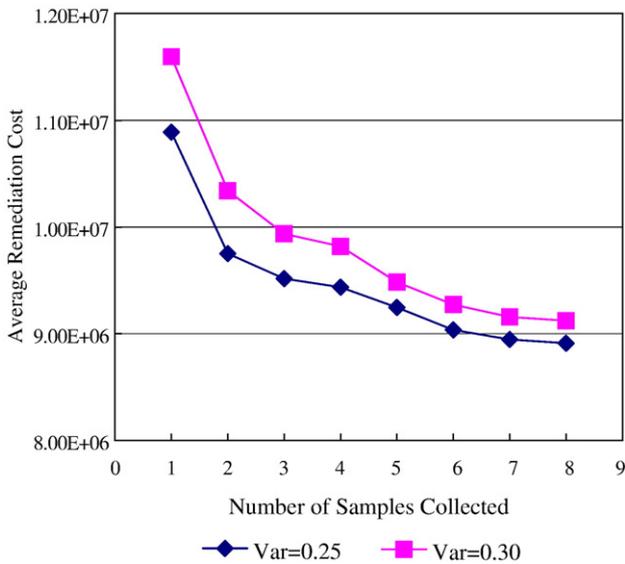


Fig. 5 – The relationships between the average remediation costs and number of samples collected under two values of variance of $\ln(K)$, 0.25 and 0.30, used in Hydro_Gen.

4.1.2. The effect of the range parameter on the value of information

Fig. 6 shows the two curves of average remediation cost with the number of samples for two values of the range parameter, 20 and 30. It indicates that the magnitude of cost decrease due to sample increase is larger at higher values of the range parameter. A higher range value means that the range over which a sampling point exerts significant influence is larger and thus sampling produces greater reduction of cost.

4.5. Expected risk and sampling numbers

Fig. 7 graphs the relationships between the expected risk at 99th percentile confidence level and the number of collected

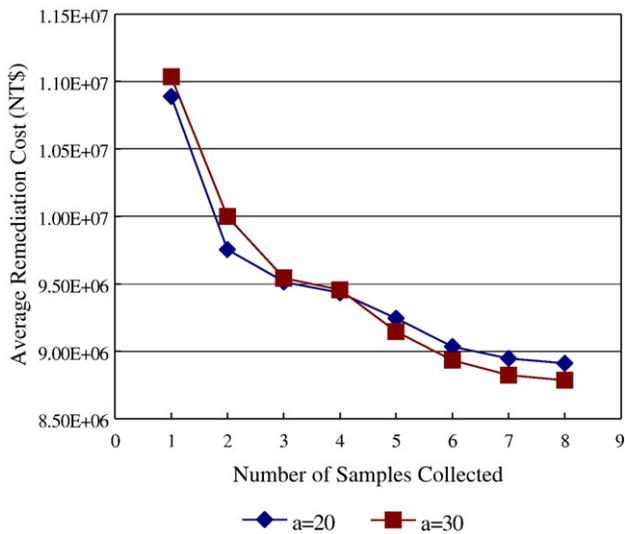


Fig. 6 – The relationships between the average remediation costs and number of samples collected under two values of the range parameter, 20 m and 30 m, used in Hydro_Gen.

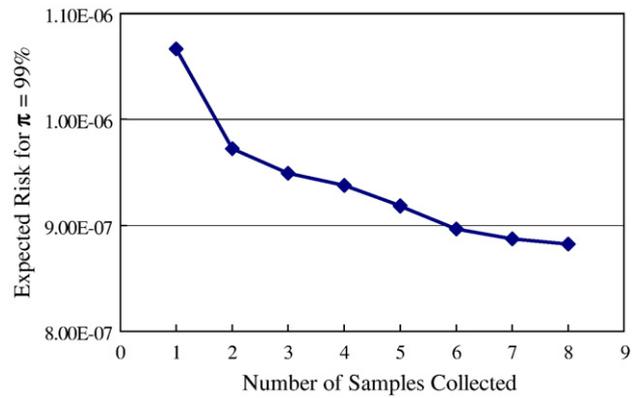


Fig. 7 – The relationships between the expected risk at 99th percentile confidence level and the number of collected samples.

samples. It is found that the 5 sampling programs’s average risk decreases with increasing numbers of samples and has decreasing marginal change. Both Figs. 5 and 7 can be used to evaluate the value of sampling information for the purpose of designing effective and efficient remediation strategies.

4.6. The effect of number of random field realizations

It takes much time to perform modeling in the multiple-realization management model. If fewer numbers of realizations produce similar outputs, then a lot of time can be saved. Fig. 8 shows the variation (represented by coefficients of variation) of the risks resulting from the N simulated random fields that went through the same 6 months without remediation for different numbers of samples. Fig. 9 shows the risk at the 99% confidence level among the N simulated fields for

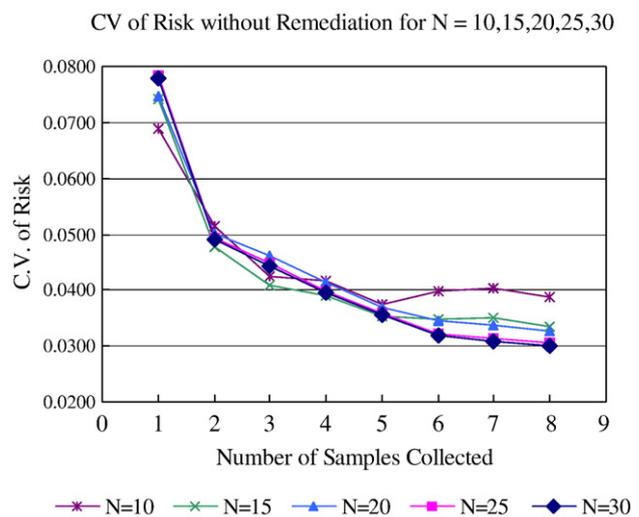


Fig. 8 – The variation (represented by coefficients of variation) of the risks resulting from the N (N=10, 15, 20, 25, and 30, respectively) simulated random fields that went through the same 6 months without remediation for different numbers of samples.

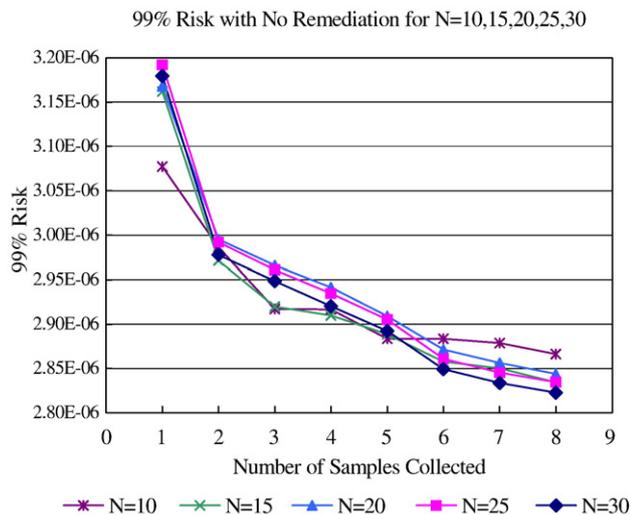


Fig. 9—The risk at the 99% confidence level among the N ($N = 10, 15, 20, 25,$ and $30,$ respectively) simulated fields for different numbers of samples.

different numbers of samples. 5 N 's are evaluated in the two figures; they are 10, 15, 20, 25, and 30. It is found that 25 and 30 simulated fields produced quite close results; when the number of simulated fields is less than 20, the estimated value of sampling would deviate to some degree.

5. Conclusion

The adapted chance-constrained multiple-realization management model that integrates random field generation, conditional simulation, and chance-constrained optimization has been demonstrated useful to provide the value of obtaining information of a contaminated site's characteristics. In the modified multiple-realization management model, the variation among the plausible fields of hydrogeological properties of the contaminated site is considered explicitly in addition to the consideration of spatial correlation structure of regionalized variables. From both microscopic and macroscopic aspects, information collection reduces uncertainty about the knowledge of the site characteristics and can be reflected by the reduced variation of random fields. When a site is better characterized, the optimized cost of remediation to meet a specified reliability level of the remediation criteria will be reduced because the range of possible sites that need to be addressed becomes narrower. Therefore the cost reduction can represent the value of uncertainty reduction through information collection. On the other hand, when the decision objective is to reduce health risk resulting from the contamination subject to a budget constraint, the better the site is characterized, the more successful the remediation strategy that is designed based on the characterization is, and the smaller the resultant risk after remediation is. Hence risk reduction can also serve as the measure of the value of sampling information. The developed methodology can be applied to various site properties that influence the distribution of contaminants and thus the result of remediation work. Site characterization and information

collection are aimed to facilitate decision making; they should be evaluated and guided in terms of the value of the effort to the decision purposes. This study provides such a methodology to relate site characterization and information collection to decision goals.

REFERENCES

- Aly AH, Peralta RC. Optimal design of aquifer cleanup systems under uncertainty using a neural network and genetic algorithm. *Water Resour Res* 1999;35:2523–32.
- Bau DA, Mayer AS. Data-worth analysis for multiobjective optimal design of pump-and-treat remediation systems. *Adv Water Resour* 2007;30:1815–30.
- Bellin A, Rubin Y. HYDRO_GEN: a spatially distributed random field generator for correlated properties. *Stoch Environ Res Risk Assess* 1996;10:253–78.
- Delhomme JP. Spatial variability and uncertainty in groundwater flow parameters: a geostatistical approach. *Water Resour Res* 1979;15:269–80.
- Geomatrix Consultants Inc. Groundwater Remediation Technical Practicability Evaluation Report. Taoyuan Site; 1998.
- Harbaugh AW, Banta ER, Hill MC, McDonald MG. MODFLOW-2000. The U.S. Geological Survey Modular Ground-Water Model—User Guide to Modularization Concepts and the Ground-Water Flow Process. U.S. Geological Survey; 2000.
- Hoeksema RJ, Kitanidis PK. Analysis of spatial structure of properties of selected aquifers. *Water Resour Res* 1985;21:563–72.
- James BR, Gorelick SM. When enough is enough: the worth of monitoring data in aquifer remediation design. *Water Resour Res* 1994;30:3499–513.
- Ma HW. Stochastic multimedia risk assessment for a site with contaminated groundwater. *Stoch Environ Res Risk Assess* 2002;16:464–78.
- Mantoglou A, Wilson JL. The Turning Bands method for simulation of random fields using line generation by a spectral method. *Water Resour Res* 1982;18:1379–97.
- Mantoglou A, Kourakos G. Optimal groundwater remediation under uncertainty using multi-objective optimization. *Water Resour Manage* 2007;21:835–47.
- McKinney DC, Lin MD. Pump-and-treat groundwater remediation system optimization. *J Water Resour Plan Manage* 1996;122:128–36.
- Meyer PD, Brill ED. A method for location wells in a groundwater monitoring network under conditions of uncertainty. *Water Resour Res* 1988;24:1277–82.
- Meyer PD, Valocchi AJ, Eheart JW. Monitoring network design to provide initial detection of groundwater contamination. *Water Resour Res* 1994;30:2647–59.
- Mitchell M. An introduction to genetic algorithms. Massachusetts Institute of Technology; 1996.
- Nicholas PC. Groundwater remediation and treatment technologies. Noyes Publication; 1997.
- Palisade Corporation. Guide to evolve: the genetic algorithm solver for Microsoft excel. Palisade Corporation; 2001.
- Reed P, Minsker B, Valocchi AJ. Cost-effective long-term groundwater monitoring design using a genetic algorithm and global mass interpolation. *Water Resour Res* 2000;36:3731–41.
- Ricciardi KL, Pinder GF, Karatzas GP. Efficient groundwater remediation system design subject to uncertainty using robust optimization. *J Water Resour Plan Manage* 2007;133:253–63.
- Rouhani S, Hall TJ. Geostatistical schemes for groundwater sampling. *J Hydrol* 1988;103:85–122.
- Storck B, Eheart JW, Valocchi AJ. A method for the optimal location of monitoring wells for detection of groundwater

- contamination in three-dimensional heterogeneous aquifers. *Water Resour Res* 1997;33:2081–8.
- Van Groenigen JW. The influence of variogram parameters on optimal sampling schemes for mapping by kriging. *Geoderma* 2000;97:223–36.
- Wagner BJ. Sampling design methods for groundwater modeling under uncertainty. *Water Resour Res* 1995;31:2581–91.
- Wagner BJ. Evaluating data worth for groundwater management under uncertainty. *J Water Resour Plan Manage* 1999;125:281–8.
- Wagner BJ, Gorelick SM. Optimal groundwater quality management under parameter uncertainty. *Water Resour Res* 1987;23:1162–74.
- Wagner BJ, Gorelick SM. Reliable aquifer remediation in the presence of spatially variable hydraulic conductivity: from data to design. *Water Resour Res* 1989;25:2211–25.
- Warner JW, Tamayo-Lara C, Khazaei E, Manghi F. Stochastic management modeling of a pump and treat system at the Rocky Mountain Arsenal near Denver, Colorado. *J Hydrol* 2006;328:523–37.
- Zheng C, Wang PP. MT3DMS: a modular three-dimensional multispecies transport model for simulation of advection, dispersion, and chemical reaction of contaminants in groundwater system; documentation and user's guide. U.S. Army Engineer Research and Development Center; 1999.