RISK-BASED IN SITU BIOREMEDIATION DESIGN

BY

JENNINGS BRYAN SMALLEY

A.B., Washington University, 1992

THESIS

**Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Environmental Science in Civil Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1998**

Urbana, Illinois

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my research advisor, Dr. Barbara Minsker, for the opportunity to work on a research project that challenged and taught me so much. I am grateful for Dr. Minsker's guidance, insight, and support while performing my research and writing my thesis.

I want to thank Jeremy Kosegi, my officemate, for his friendship and discussions about our research. The long days and nights in the office in front of the workstation were much more bearable and productive as a result. I would also like to thank graduate student Patrick Reed for his assistance in formulating and quantifying some of the remediation cost equation components.

I am grateful to Dr. David Goldberg and Dr. Brad Miller for their discussions and advice regarding the theory and application of noisy genetic algorithms. I also want to thank Dr. Tim Ellsworth for his help with some of the geostatistical aspects of this research.

Finally, I would like to thank my parents for showing me both the academic and practical importance of an education. They, and my entire family, have always encouraged and supported me throughout all of my academic pursuits.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Introduction and Objectives

Due to the adverse impacts of petroleum hydrocarbons on human health and the environment, considerable resources have been expended to restore sites contaminated with petroleum constituents. In the past, cleanup goals were often established without regard to risk, mandating remediation of groundwater to background or non-detection levels, to maximum contaminant levels, or to some level of total petroleum hydrocarbons. Such practices have produced goals that are often difficult or impossible to achieve and have made site restoration prohibitively expensive.

Risk-based corrective action (RBCA) is rapidly becoming an accepted approach to remediating contaminated sites. Under a RBCA approach, the risks to human health and the environment associated with a contaminated site are evaluated and appropriate corrective measures are taken as needed to reduce risk to acceptable levels. Numerous states have implemented or are in the process of implementing RBCA standards for petroleum-release sites based on procedures developed in 1995 by the American Society for Testing and Materials [*ASTM*, 1995]. ASTM describes RBCA as a "tiered" approach to risk assessment with movement from a lower to a higher tier only when necessary. Within each tier, risk assessment and remedial action are appropriately tailored to the extent of available site assessment data. A Tier 1 evaluation involves comparing conservative, generic screening-level concentrations, or Risk-Based Screening Levels (RBSL's), to site conditions. Higher tiers (Tiers 2 or 3) involve a greater degree of sophistication and expense for data collection and modeling but may allow overall cost savings because site-specific target levels (SSTL's) are established as remediation goals. Progression to a higher tier is warranted when the assumptions made to develop the RBSL's are inappropriate relative to site conditions or when the expenditures associated with

higher tier evaluation and subsequent corrective action are lower than the preceding tier evaluation and remediation costs.

Tier 2 and Tier 3 evaluations will likely involve the use of predictive models to assess risks and to evaluate the effectiveness of candidate remedial designs. Recently, a significant amount of work has been performed to research and develop simulation models that integrate contaminant fate and transport and exposure and risk assessment. Multimedia models such as MMSOILS [*U.S. EPA*, 1992] and MEPAS [*Droppo et al.*, 1989] are comprehensive fate and transport and exposure and risk assessment models that consider multiple processes, exposure scenarios and routes, and contamination scenarios. *Labieniec et al.* [1996a, 1996b] developed a less comprehensive but more computationally efficient analytical multimedia model that can be run stochastically, using probability density functions for certain site- and contaminant-related parameters. Another recent model described in *Pelmulder et al.* [1996] simulates contaminant fate and transport in groundwater and assesses exposure and risk resulting from the use of contaminated well water in the home and for irrigation.

In *Andricevic and Cvetkovic* [1996], a methodology was described for predicting the risk level for individuals ingesting groundwater contaminated with radioactive material. Their stochastic approach used first-order first- and second-moment analysis to account for parameter uncertainty and variability associated with source term characterization, contaminant fate and transport, and risk assessment. *Maxwell et al.* [1998] developed a stochastic model that combines contaminant fate and transport simulation and multiple pathway exposure and risk assessment. Their model assumes that contaminated groundwater is withdrawn at multiple wells that supply water to residents, and the authors consider the effects on exposure and risk assessment of variations in the configuration of the water distribution system. The model also

incorporates the effects of model parameter uncertainty and variability by using numerous unconditional hydraulic conductivity realizations and by randomly sampling from a number of exposure parameter distributions. The work of *Davis et al.* [1997] presented a risk-based approach to soil remediation modeling that accounts for heterogeneity in contaminant distributions and site-specific vadose and saturated zone parameters. The authors used their model to iteratively select soil parcels for excavation until risk-based standards were met at hypothetical receptor wells.

While these simulation models can be helpful for RBCA design, groundwater management models are generally more useful for design than simulation models alone. Management models combine optimization procedures with simulation models in order to develop remedial designs that are both cost-effective and reliable. Such models offer two advantages over simulation models alone: (1) given a virtually unlimited number of possible design options (such as well locations and pumping rates), it is unlikely that a least-cost and reliable design will be found using trial-and-error simulation alone; and (2) optimization allows accurate and informed comparisons to be made between remediation alternatives (e.g. enhanced in situ bioremediation vs. pump and treat) because both designs are optimal for the given site conditions. Groundwater management models developed to date have utilized a wide variety of optimization techniques, remediation strategies, and definitions of what constitutes an "optimal design."

These management models have used traditional gradient-based nonlinear programming methods for pump-and-treat remedial design [*McKinney and Lin*, 1995, 1996], cutting plane optimization techniques for pump-and-treat design [*Karatzas and Pinder*, 1993, 1996], and combinatorial optimization methods such as genetic algorithms and simulated annealing.

Genetic algorithms have been used for pump-and-treat management models that have contaminant constraints [*McKinney and Lin*, 1994] and for reliable plume containment design [*Ritzel et al.*, 1994]. Simulated annealing has also been applied to contaminant-constrained groundwater management problems [*Dougherty and Marryott*, 1991; *Marryott et al.*, 1993].

In other groundwater quality management models, the definition of "optimal design" has been extended beyond cost-effectiveness to include design reliability. These models quantify and incorporate uncertainty resulting from parameter estimation and from the spatial variability of aquifer properties into the optimization process. These stochastic approaches have included parameter estimation and first-order analyses for groundwater contaminant capture design [*Tiedeman and Gorelick*, 1993] and for pump-and-treat remediation design [*Wagner and Gorelick*, 1987]. Other approaches have accounted for the spatial variability of hydraulic conductivity. *Wagner and Gorelick* [1989] used multiple conditional realization methods for pump-and-treat design, while *Morgan et al.* [1993] and *Ritzel et al.* [1994] incorporated multiple realizations into their contaminant plume containment design procedures. Finally, groundwater management models have been developed that include time-dependent design variables, allowing additional cost savings relative to time-invariant approaches. These dynamic optimal control methodologies have included strategies for pump-and-treat remediation [*Culver and Shoemaker*, 1992; *Chang et al.*, 1992] and, more recently, for in-situ bioremediation of groundwater [*Minsker and Shoemaker*, 1998].

Although these management models have been shown to be effective for reducing costs associated with traditional remediation approaches, they have not yet been applied within the RBCA framework. In this thesis, the development of a risk-based groundwater management model is described. The model combines contaminant fate and transport simulation models and

health risk assessment procedures with an optimal search algorithm, simultaneously predicting human health risk and proposing cost-effective strategies for reducing risk to an acceptable level. A key aspect of this coupled management model is the ability to account for the uncertainty and variability in parameters associated with groundwater simulation modeling and the risk assessment process. Accounting for parameter uncertainty and variability will support the development of risk reduction strategies that are both cost-effective and sufficiently but not overly conservative. A noisy genetic algorithm (GA) is used to incorporate parameter uncertainty and variability within the search for optimal remedial designs. Noisy GAs are relatively new techniques within the evolutionary computing field; in this work, their potential for application within groundwater management models is demonstrated.

Evaluation of natural attenuation and in situ bioremediation strategies will likely be an important part of RBCA, particularly for petroleum releases. This work focuses on in situ bioremediation as the first remediation technology to be addressed, although future versions of the model will also be capable of evaluating any combination of pump-and-treat, natural attenuation, and in situ bioremediation. In order to include this capability, the model allows cleanup times to be variable, depending on the well locations and pumping rates selected, unlike most previous groundwater management models which allow only fixed cleanup lengths. Following a description of the management model components, implementation of the model is demonstrated by applying it to a realistic remediation scenario.

2. Management Model

2.1 Overview

The risk-based groundwater management model is a combination of several components, which are described in detail in Sections 2.2-2.7. Figure 1 depicts each component and illustrates how the management model is solved. First, the model user specifies a number of initial trial designs. The numerical groundwater flow and contaminant fate and transport model (A), described in Section 2.3, is used to evaluate the effectiveness of the trial designs by simulating contaminant concentrations in and near the source area over time. The analytical fate and transport model (B), described in Section 2.4, then uses numerical model predictions of contaminant concentrations in the source area to predict steady-state concentrations at drinking water wells significantly further downgradient from the source. These exposure concentrations are provided as inputs to the exposure and risk assessment model (C), described in Section 2.5. (Section 2.4 explains why steady-state concentrations from the analytical model are used in lieu of transient concentrations from the numerical model.) The exposure and risk model combines exposure assessment procedures and toxicological information in order to estimate human health risk. Using the results from the simulation models, the trial designs are evaluated in terms of the objective function and constraints (D), given in Section 2.2. The objective function and constraints represent the goals of the management model, which are to minimize the remedial cost and to achieve groundwater quality, gauged in terms of human health risk, while also conforming to other necessary design restrictions. Though not pictured here, an additional goal of the model is to provide a degree of design reliability by incorporating uncertainty and variability in model parameters through Monte Carlo evaluations of the objective function and constraints for each trial design (described in Section 2.6). Given the results of these

evaluations, a noisy genetic algorithm (E) performs procedures analogous to natural selection in order to generate a set of new trial designs as described in Section 2.7. The above procedure is repeated until the trial designs are acceptable under specified stopping criteria (discussed in more detail in Section 2.7).

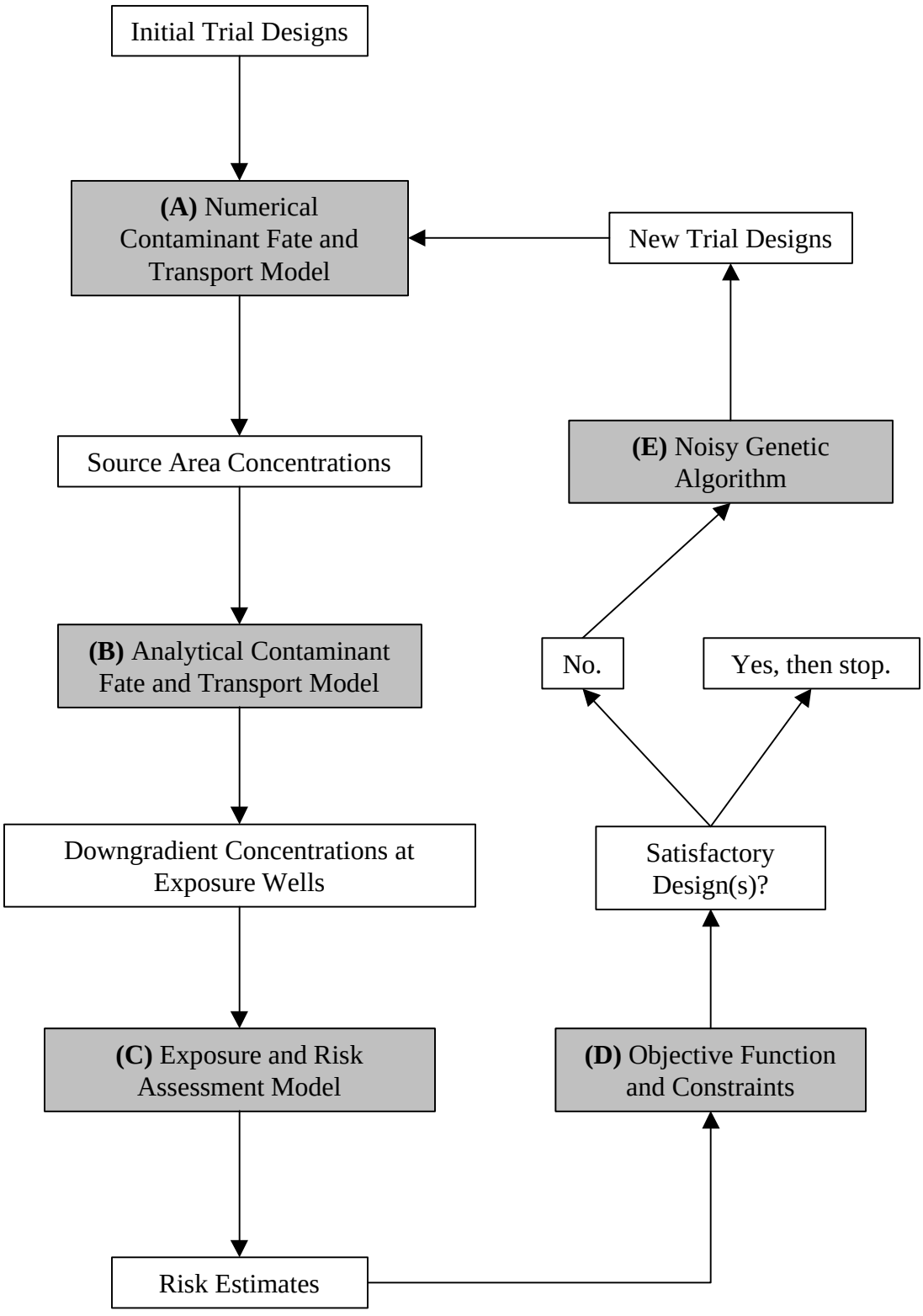

Figure 1. Management Model Components

2.2 Objective Function and Constraints

The objective function (1) represents one of the main goals of the management model: minimizing the total present value cost of a bioremediation design. C_{TOT} includes capital and operating costs for remediation wells, C_{REM}, site monitoring costs, C_{MON}, as well as additional capital and operating costs associated with the remedial system, C_{SYST}. These individual cost components are shown in Equations (2)-(4),

$$Min\ C_{TOT} = C_{REM} + C_{MON} + C_{SYST} \tag{1}$$

$$C_{REM} = \sum_{i=1}^{N} (c_i^{cap,rem} x_i + c_i^{op,rem} \mid u_{i,l} \mid t_{cu}) \tag{2}$$

$$C_{MON} = c^{capsite,mon} + \sum_{j=1}^{M} c_j^{capwell,mon} + \left[c^{samp,mon} + \sum_{j=1}^{M} c_j^{ana,mon} \right] N_{mon} \tag{3}$$

$$C_{SYST} = \left[c^{cap,syst} + \left(\sum_{i=1}^{N} u_{i,l} c^{muni,syst} + c^{op,syst} \right) t_{cu} + c^{main,syst} Y_{cu} \right] X_{syst} \tag{4}$$

where N is the total number of remediation wells (injection and extraction), $c_i^{cap,rem}$ is the capital cost associated with the installation of each remediation well i (\$), x_i is an installation decision variable for remediation well i, $c_i^{op,rem}$ is the operating cost for remediation well i (\$/m^3), $u_{i,l}$ is a decision variable for the pumping rate for well i ($u_{i,l}$ is positive for injection and negative for extraction) (m^3/day), l is a decision subscript representing the location of each remediation well,

t_{cu} is the length of the remediation period [a calculated variable where t_{cu} is a function of $u_{i,l}$ and l] (days), $c^{capsite,mon}$ is a cost associated with monitoring well installation at the site ($), M is the total number of monitoring wells [fixed by the user], $c_j^{capwell,mon}$ is the capital cost associated with the installation of each monitoring well j ($), $c^{samp,mon}$ is the cost to collect samples at the site ($), $c_j^{ana,mon}$ is the cost to analyze samples at each monitoring well j ($), N_{mon} is the number of monitoring periods within the remediation period, $c^{cap,\,syst}$ includes additional capital costs associated with the remedial system ($), $c^{muni,syst}$ is the cost for municipally-supplied water ($/m^3), $c^{op,syst}$ includes costs for additional system operation and management ($/day), $c^{main,syst}$ is the cost for equipment maintenance ($), Y_{cu} is the number of whole years in the remediation period, and X_{syst} is a remedial system installation and operation indicator variable which has a value of 1 if any remediation wells are installed and a value of zero otherwise.

The optimization algorithm searches for sets of pumping rates $u_{i,l}$ and well locations l (given a specified number of potential wells and locations) that minimize the total remedial cost [Equation (1)], subject to the following constraint equations:

$$Risk_{t,k}^{TOTAL} = Risk_{t,k}^{w} + Risk_{t,k}^{shw} + Risk_{t,k}^{nc} \leq TR \qquad \forall t, \forall k \tag{5}$$

$$u_{min} \leq |u_{i,l}| \leq u_{max} \qquad \forall i \tag{6}$$

$$l \in \Omega_L \qquad \forall l \tag{7}$$

$$h_{min,l} \leq h_{i,l} \leq h_{max,l} \qquad \forall i \text{ at each } l \tag{8}$$

$$\sum_{i=1}^{N} u_{i,l} \geq 0 \tag{9}$$

$$x_i = \begin{cases} 0 & \text{then } u_{i,l} = 0 \\ 1 & \text{then } u_{i,l} = u_{i,l} \end{cases} \qquad \forall i \tag{10}$$

where $Risk_{t,k}^{TOTAL}$ is the total individual lifetime health risk at time t and exposure location k; $Risk_{t,k}^{w}$, $Risk_{t,k}^{shw}$, and $Risk_{t,k}^{nc}$ are the risk due to ingestion of contaminated drinking water, inhalation of volatiles from contaminated water due to showering, and inhalation of volatiles from contaminated water due to other non-consumptive uses, respectively, from concentrations at time t and exposure location k (see Sections 2.4 and 2.5 for details on how these are calculated); k represents the downgradient location of a well for actual or potential human consumption and use; TR is the target total individual lifetime health risk; u_{min} and u_{max} are the minimum and maximum absolute pumping rate for a given remediation well (m^3/day); Ω_L is a set of potential remediation well locations; and $h_{i,l}$, $h_{min,l}$, and $h_{max,l}$ are the calculated hydraulic head for remediation well i (m), the minimum hydraulic head (m), and the maximum hydraulic head (m) at remediation well location l, respectively.

Equations (5)-(10) represent: an upper bound on the total individual lifetime health risk TR at all times t within the remediation period and at the end of the remediation period t_{cu} (5), lower and upper bounds on the remediation well pumping rates (6), a restriction on potential remediation well locations (7), lower and upper bounds on the hydraulic head at remediation

well locations (8), a restriction that the total rate of extraction cannot be greater than that of injection [because any extracted water is re-injected due to the absence of above-ground treatment] (9), and pumping rate specifications depending on the value of the remediation well installation decision variable (10).

2.3 Numerical Model

The function of the numerical simulation component of the coupled management model is to evaluate the effectiveness of a given bioremediation design. Optimal design of an in situ bioremediation strategy is assumed to involve determining the locations and pumping rates for injection and extraction wells. For petroleum hydrocarbons, injection wells are used primarily to stimulate microbial growth and to accelerate degradation of the contaminant, known as the electron donor or primary substrate, by injecting an electron acceptor, such as oxygen. Degradation of key petroleum constituents, benzene, toluene, ethylbenzene and xylene (BTEX), has been most successful under aerobic conditions with oxygen as the electron acceptor [for reference, see *Hutchins*, 1991].

The numerical simulation model is a two-dimensional, depth-averaged finite-element flow, transport, and biodegradation model called Bio2D [*Taylor*, 1993]. The model predicts hydraulic heads and substrate (contaminant), oxygen, and microbial biomass concentrations in a confined aquifer that would result from pumping rates and well locations selected by the optimization component of the management model. There are four governing equations that characterize this model. The first describes steady-state groundwater flow and is solved once for a given set of pumping locations and rates. The three remaining equations describe time-dependent changes in substrate (contaminant), oxygen, and microbial biomass concentrations.

Changes in concentrations over time result from the processes of advection, dispersion, adsorption and biodegradation. The rates of substrate (contaminant) and oxygen degradation and of microbial biomass growth are modeled using the Haldane variant of the Monod equation [for reference, see *Gaudy and Gaudy*, 1980]. This equation includes a term that allows inhibition of microbial growth to occur at high substrate (contaminant) concentrations. Bio2D uses a Galerkin finite-element method in space and a variably weighted, backward-difference approximation in time to find solutions to the four governing equations. More details on Bio2D, including the governing equations, are given in Appendix A. A discussion of assumptions and limitations associated with the use of Bio2D is given in Appendix A and *Minsker and Shoemaker* [1998].

2.4 Analytical Model

A steady-state analytical model is used to predict groundwater concentrations further downgradient from the contaminated site at exposure wells, using numerically modeled source area concentrations as inputs. An analytical model is implemented for predicting groundwater exposure concentrations for two reasons: (1) the analytical model simulates the advection, dispersion, adsorption, and degradation of contaminant over greater periods of time and distance than would be possible within the framework of a computationally demanding management model, and (2) the protocol and practice for using exposure concentrations from a transient numerical model for estimating total lifetime health risk have not been established. RBCA guidance [*ASTM*, 1995] promotes the use of a constant groundwater concentration over the duration of exposure when developing most remedial action target levels. In addition, *U.S. EPA* [1989] advocates the use of the average concentration that is contacted over the exposure period

for long-term contact, and *U.S. EPA* [1988] states that long-term average "concentration values are obtained from steady-state methods."

The analytical model is similar to the one given in RBCA guidance [*ASTM*, 1995] and reference should be made to *Domenico and Schwartz* [1990] for more detailed information. Although downgradient concentrations are only predicted along the centerline in the x-direction, the model accounts for the effects of transverse dispersion as well as decreases in concentration due to longitudinal dispersion, degradation, and adsorption. The analytical model is given in Equation (11),

$$C_{t,k}^{w} = C_{t}^{source} \exp\left\{\frac{x_k}{2a_x}\left[1 - \sqrt{\left(1 + \frac{4 l_b a_x}{v}\right)}\right]\right\} erf\left[\frac{S_{wid}}{4\sqrt{a_y x_k}}\right] \qquad (11)$$

where C_t^{source} is a numerically-modeled source area contaminant concentration in groundwater at time t [mg/L], x_k is the distance, along the centerline, from the location of a source area concentration to a human use well at location k [m], a_x is the longitudinal dispersivity [m], l_b is the first-order degradation constant [1/day], v is the average linear groundwater velocity [m/day], S_{wid} is the width of the source area contaminant plume [m], a_y is the transverse dispersivity [m], and $C_{t,k}^{w}$ is a steady-state contaminant concentration in groundwater predicted along the centerline at location k based on a source area concentration at time t [mg/L].

The C_t^{source} and S_{wid} terms are derived from a monitoring network established in the numerically modeled source area. Likewise, v is based on numerical model predictions of steady-state groundwater velocities in the source area without pumping. The derivation of the

values for these and the other analytical parameters in Equation (11) are described in more detail in Section 3.

2.5 Exposure and Risk Assessment Model

This section details the methodology for the assessment of human exposure and for the computation of human health risk. Before discussing exposure and risk assessment procedures in detail, though, it is necessary to discuss the meaning of human health risk within the RBCA framework and within the context of this research. Risk-based standards, or remedial action target levels, are based on measures that gauge the effect of human exposure to a given compound. RBCA guidance has been established for human exposure to both carcinogenic and non-carcinogenic substances. For non-carcinogens, this gauge is the "hazard quotient", whereas the one for carcinogens is simply termed "risk". Although the effects of non-carcinogenic exposure can be readily integrated into the management model, the current formulation considers only carcinogenic compounds.

Risk is a unitless number defined as the "upper bound lifetime probability of an individual developing cancer as a result of exposure to a particular level of a potential carcinogen" [*U.S. EPA*, 1989]. For example, a risk of 10^{-6} represents an increased probability of one-in-one-million. Because conservative procedures and estimates are used for assessments, the actual risk to an exposed population is considerably lower. There is also the implicit assumption that the specified exposure scenario occurs. In practice, estimates of risk function as a means to separate low-risk situations from those that deserve regulatory and/or remedial action. Although there is considerable debate regarding the amount of allowable risk [see *ASTM*, 1995], values in the range of 10^{-4} to 10^{-6} are usually considered acceptable.

The exposure and risk assessment model combines predicted groundwater contaminant concentrations with relevant exposure parameters and toxicological information to determine the total individual lifetime health risk [$Risk_{t,k}^{TOTAL}$ in Equation (5)]. This research assumes an off-site residential exposure scenario: groundwater is obtained from a well some significant distance downgradient from the contaminated site and is used in a nearby-occupied residence(s).

Within the residential scenario, three exposure routes are defined: ingestion of contaminated drinking water ($Risk_{t,k}^{w}$), inhalation of volatiles from contaminated water due to showering ($Risk_{t,k}^{shw}$), and inhalation of volatiles from contaminated water due to other non-consumptive uses such as dish and clothes washing and cooking ($Risk_{t,k}^{nc}$). *McKone* [1987] asserts that the exposure to volatiles in drinking water by inhalation during non-consumptive use may be as great or greater than that due to direct water ingestion. Additional exposure routes are possible but are not considered. Specifically, dermal absorption of contaminated water during showering is neglected because *Schaum et al.* [1994] concluded that this route is insignificant relative to ingestion and inhalation for most contaminants.

The risk associated with each exposure route is computed by multiplying the lifetime average daily-absorbed dose of contaminant by the appropriate cancer potency slope factor:

$$Risk_{t,k}^{w} = D_{t,k}^{w} \cdot CPS_o \tag{12}$$

$$Risk_{t,k}^{shw} = D_{t,k}^{shw} \cdot CPS_i \tag{13}$$

$$Risk_{t,k}^{nc} = D_{t,k}^{nc} \cdot CPS_i \tag{14}$$

where $D_{t,k}^{w}$, $D_{t,k}^{shw}$, and $D_{t,k}^{nc}$ are the absorbed dose of contaminant due to ingestion of contaminated water, inhalation of volatiles due to showering, and inhalation of volatiles due to other uses, respectively (mg/day-kg); and CPS_o and CPS_i are the oral and inhaled cancer potency slope factors, respectively (mg/day-kg)$^{-1}$. Slope factors are chemical-specific estimates of risk per unit dose of a potential carcinogen. Because (carcinogenic) risk is additive, the total risk, $Risk_{t,k}^{TOTAL}$, is the sum of the risk from each exposure route as shown in Equation (5).

$D_{t,k}^{w}$, $D_{t,k}^{shw}$, and $D_{t,k}^{nc}$ in Equations (12)-(14) are calculated as follows:

$$D_{t,k}^{w} = \frac{C_{t,k}^{w} \cdot IR_{w,ad} \cdot EF \cdot ED_{ad}}{BW_{ad} \cdot AT} + \frac{C_{t,k}^{w} \cdot IR_{w,ch} \cdot EF \cdot ED_{ch}}{BW_{ch} \cdot AT} \quad (15)$$

$$D_{t,k}^{shw} = \frac{C_{t,k}^{a,shw} \cdot IR_{a,ad} \cdot (SD + BD) \cdot EF \cdot ED_{ad}}{BW_{ad} \cdot AT} +$$

$$\frac{C_{t,k}^{a,shw} \cdot IR_{a,ch} \cdot (SD + BD) \cdot EF \cdot ED_{ch}}{BW_{ch} \cdot AT} \quad (16)$$

$$D_{t,k}^{nc} = \frac{C_{t,k}^{a,nc} \cdot IR_{a,ad} \cdot ID \cdot EF \cdot ED_{ad}}{BW_{ad} \cdot AT} +$$

$$\frac{C_{t,k}^{a,nc} \cdot IR_{a,ch} \cdot ID \cdot EF \cdot ED_{ch}}{BW_{ch} \cdot AT} \quad (17)$$

where $C_{t,k}^{w}$ is the contaminant concentration in groundwater at a well extracted for human consumption and use (mg/L) [as computed by the analytical model in Equation (11)], $IR_{w,ad}$ is the adult drinking water ingestion rate (L/day), $IR_{w,ch}$ is the child drinking water ingestion rate (L/day), *EF* is the residential exposure frequency (days/year), ED_{ad} is the adult exposure duration (years), ED_{ch} is the child exposure duration (years), BW_{ad} is the adult body weight (kg), BW_{ch} is the child body weight (kg), *AT* is the averaging time (days), $C_{t,k}^{a,shw}$ is the contaminant concentration in shower/bathroom air (mg/m^3), $IR_{a,ad}$ is the adult air inhalation rate (m^3/day), $IR_{a,ch}$ is the child air inhalation rate (m^3/day), *SD* is the shower duration (days/day), *BD* is the time spent in the bathroom after shower (days/day), $C_{t,k}^{a,nc}$ is the contaminant concentration in household air (m^3/day), and *ID* is the time spent indoors (days/day).

Equations (15)-(17) are taken from EPA guidance [*U.S. EPA*, 1989] and are in accordance with those in ASTM's RBCA guidance [*ASTM*, 1995]. Note that the absorbed dose for each route is the sum of both adult and child exposure. The methods used in this research to calculate the contaminant concentration in shower/bathroom air ($C_{t,k}^{a,shw}$) and indoor air ($C_{t,k}^{a,nc}$), given the contaminant concentration in extracted groundwater ($C_{t,k}^{w}$), are taken from *Schaum et al.* [1994] and are discussed in detail in Appendix A.

2.6 Uncertainty, Variability, and Reliability

Current regulatory guidance under the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) and the Resource Conservation and Recovery Act (RCRA) requires only single-point estimates of risk for contaminants at relevant sites. These estimates are based on a combination of values for exposure parameters that result in the

"reasonable maximum exposure" (RME) for a given exposure route. For example, U.S. EPA's default exposure assumptions use numerous upper bound (maximum and 95th percentile) estimates for parameter values. The U.S. EPA intended for its human health risk assessment procedures to produce protective, rather than best or most likely estimates of risk [*U.S. EPA*, 1988, 1989].

The U.S. EPA recognizes that there are numerous factors within the assessment process that contribute to uncertainty in risk estimates and encourages performing analyses (e.g., Monte Carlo simulation) to track the magnification of uncertainty through the assessment stages (U.S. EPA 1988, 1989). Such analyses produce "multiple descriptors of risk." Even so, such analyses cannot be used " as a basis for the RME" [*U.S. EPA*, 1989] and are only "an optional addition to, not a substitute for, current risk assessment methods" [*U.S. EPA*, 1994]. Single-point, RME estimates of risk may be suitable for screening purposes but may not be appropriate remedial targets; i.e., risk over-estimation would lead to unnecessary remedial effort and cost. For example, *Kangas* [1996] demonstrates that an RME point estimate can produce a risk value that is greater than the 100th percentile of the Monte Carlo simulation distribution. Even in the absence of such over-estimation, "multiple descriptors of risk", in the very least, provide decision-makers with more complete information and allow for tradeoffs to be made between cleanup cost, degree of acceptable risk, and probability of risk standard violation.

One of the more difficult problems associated with the simulation-optimization approach to groundwater quality management is insuring that a remedial design with a low probability of human health risk standard violation is obtained. Achievement of a reliable design requires identifying, quantifying, and incorporating relevant sources of model parameter uncertainty and variability into the optimization process. The *U.S. EPA* [1989] identifies several sources of

uncertainty or variability: (1) definition of physical setting (land uses, exposure pathways, substances of concern), (2) model applicability and assumptions, and (3) transport, fate and exposure parameter values. This research focuses only on uncertainty in the ultimate risk estimate that is derived from the third source, model parameter uncertainty and variability. These parameters can include, but are not limited to, physical, chemical, and biodegradative parameters associated with the groundwater simulation model (Equations (20)-(23) in Appendix A) and exposure model parameters used to calculate lifetime average daily absorbed dose, as shown in Equations (15)-(17). The uncertainty and variability of these parameters is discussed in more detail in Sections 2.6.1 and 2.6.2 below. It is important to distinguish between model parameters that are uncertain (such as spatially distributed groundwater parameters) and those that are variable (such as exposure model parameters that describe an individual's physical characteristics, behaviors, and surroundings). The term "uncertainty" is used throughout this work to describe both the effects of model parameter uncertainty and variability on the ultimate risk estimate. These effects are incorporated into the management model using the methodology described in Section 2.6.3 below.

2.6.1 Groundwater Parameters

Much uncertainty in groundwater modeling results from a lack of knowledge regarding the natural spatial variability of porous media parameters, such as hydraulic conductivity. Under natural conditions, the heterogeneity of hydraulic conductivity can be a key factor in determining the transport and ultimate fate of groundwater contaminants. During bioremediation, the heterogeneity influences the rate and ultimate extent of contaminant biodegradation by controlling the ability of injection wells to deliver oxygen and/or nutrients to the contaminant. Because sampling to determine such parameters is expensive and often difficult, it is impossible

to fully characterize the spatial variability. Even the best techniques, properly implemented, will produce uncertainty in the estimates of hydraulic conductivity. Therefore, the accuracy of groundwater simulation model predictions of contaminant concentrations (inputs to the exposure and risk assessment model) over space and time depends upon appropriate characterization and quantification of the uncertainty.

In order to represent the uncertainty associated with the spatial variability of hydraulic conductivity, a stochastic approach to groundwater flow and contaminant transport is implemented. Conditional simulation is used to generate multiple realizations of the hydraulic conductivity field that (1) are statistically similar to the actual site and (2) respect known information about the actual site distribution gathered through sampling. Multiple conditional realizations can be used to investigate the uncertainty of the spatial distribution of hydraulic conductivity as well as the resulting uncertainty of predictions of contaminant concentrations over space and time. The methodology used to generate conditional realizations is described in greater detail in Appendix C. Also, *Journel and Huijbregts* [1978], *Marsily* [1986] and *Gelhar* [1993] are valuable references for more information on the methodology used.

2.6.2 Exposure Parameters

Parameters in the equations used to calculate lifetime average daily-absorbed dose are quantified to represent human behavior, physical characteristics, and living environments. Any exposed population is diverse, however, and the variability associated with numerous components of the dose equations leads to uncertainty in the calculation of dose and, therefore, in the ultimate estimate of risk. As with hydraulic conductivity, a probabilistic approach is taken in order to characterize and quantify the variability attributed to parameters of the dose equations. The probabilistic assessment of risk developed here is not intended to be exhaustive.

Based on previous research [*Roseberry and Burmaster*, 1992; *Israeli and Nelson*, 1992; *Brainard and Burmaster*, 1992; *McKone and Bogen*, 1991; *ICRP*, 1975; *James and Knuiman*, 1987], ten of the exposure parameters in Equations (15)-(17) and (27)-(28) of Appendix B are assumed to be variable and defined by probability density functions: adult and child drinking water ingestion rates ($IR_{w,ad}$ and $IR_{w,ch}$), adult and child exposure durations (ED_{ad} and ED_{ch}), adult body weight (BW_{ad}), adult air inhalation rate ($IR_{a,ad}$), shower duration (SD), time spent in the bathroom after shower (BD), shower flow rate (SR_w), and total water use rate in house (HR_w). All ten of the variable parameters are assumed to be lognormally distributed and fully characterized by the mean and variance. The correlation between these exposure parameters is not accounted for in the calculations of dose even though correlation among some parameters is possible. *Smith et al.* [1992] have shown that weak correlations among parameters do not significantly affect the resulting probabilistic estimates.

2.6.3 Incorporating Parameter Uncertainty and Variability

The final step towards achieving a reliable design is incorporating the probabilistic information that characterizes the identified uncertain and variable model parameters into the optimization process. Using this information, it is possible to define multiple sample sets where each set consists of a single sample drawn randomly from the pool of generated conditional realizations and from each variable exposure model parameter distribution. The optimization algorithm, then, is required to produce a design (remediation well pumping locations and rates) that is valid for a significant number of sample sets. The reliability of the design produced by the optimization process can be evaluated by applying the design to an even greater number of sets. The reliability of a given design can be viewed as the proportion of sample sets for which

successful remediation (no constraint violations) is achieved. This general methodology is discussed in greater detail in Section 2.7.

2.7 Noisy Genetic Algorithms

2.7.1 Genetic Algorithm Theory

Genetic algorithms are a class of combinatorial methods used to solve complex optimization problems and are useful for problems that have nonconvex regions and discontinuities. GAs search for optimal solutions or designs using procedures analogous to natural selection or evolution. For GA's, the search procedures involve the evaluation of a set of trial designs and the propagation of only the "fittest" designs to future generations. *Holland* [1975] performed the initial work on GAs at the University of Michigan; a more recent text by *Goldberg* [1989] provides a comprehensive introduction of GA theory, operation, and research.

To implement a GA to solve an optimization problem, decision variables are coded as strings of binary digits or real numbers. In this management model, a binary coding is used. Individual binary strings of decision variables are then concatenated to create a longer string or "chromosome" which represents a single trial design or solution. The precision of a given decision variable representation is a function of the number of binary digits used. A discussion of decision variable precision as well as the advantages and disadvantages of binary- and real-coded GAs is not provided here but can be found elsewhere [*Goldberg*, 1989; *McKinney and Lin*, 1994]. Trial designs or chromosomes are grouped together to form a "population" which represents a single "generation."

To search for an optimal solution, the GA randomly generates an initial population of binary chromosomes or trial designs. The search procedure then involves decoding each binary

chromosome into a real-numbered trial design and evaluating its performance using a groundwater simulation model. Following simulation, the "fitness" of each trial design can be determined by calculating the objective function value and evaluating the constraint violations, if any, for each design. In order to enforce constraints, a penalty method is used. Here, the amount of each constraint violation is multiplied by a penalty coefficient or weight to determine the amount of the penalty. The penalties for all constraint violations are summed and added to the objective function value. Because the goal of this management model is cost minimization [see Equation (1) in Section 2.2], the fittest design is the one with the lowest combination of objective function value and constraint violation penalties or, in other words, the highest fitness function evaluation.

Once the fitness of each trial design or chromosome has been determined, the operations of selection (reproduction), crossover (mating), and mutation are used to generate a new population of trial designs. The operation of selection occurs first. In this management model, "binary tournament selection" is used to determine which trial designs become members of the mating pool [see *Ritzel et al.*, 1994 and *McKinney and Lin*, 1994 for further discussion of tournament selection in groundwater management models]. Here, each chromosome can participate in an unlimited number of tournaments until the mating pool equals the size of the population.

The next operation is crossover. The crossover procedure implemented here involves random mating and a single, randomly chosen crossover point. Crossover is performed on each mated pair of chromosomes with a probability of crossover, P_c, to form a new population [see *Ritzel et al.*, 1994 and *Goldberg*, 1989 for more on the operation of crossover]. The final operation is mutation. For a binary-coded GA, each bit of each chromosome resulting from

crossover is "flipped" with a probability of mutation, P_m. Good GA performance is usually achieved with a high crossover probability and a low mutation probability (P_m = 1/population size) [*De Jong*, 1975]. The process of fitness value determination using the simulation model and new population formulation by the three operations described above continues until convergence or a maximum number of generations is reached.

2.7.2 Noisy Genetic Algorithm Theory

Noisy genetic algorithms are simply ordinary genetic algorithms that operate in noisy environments. The "noise" that exists in certain environments can be defined as any factor that hinders the accurate evaluation of the fitness of a given trial design or chromosome [*Miller*, 1997]. These factors can include the use of approximate fitness functions, either by necessity when there is no known accurate fitness function or due to a desire for increased computational speed; the use of noisy data; knowledge uncertainty; and sampling and human error. To illustrate a case of inaccurate fitness evaluation, consider a trial design that has a high fitness value under certain conditions, but when evaluated under a different set of conditions, has a poor fitness value. It is unknown which evaluation represents the "true" fitness of this trial design. Because noise produces inaccurate or noisy fitness function evaluations, these fitness functions are termed noisy fitness functions.

As discussed in *Miller* [1997], a type of noisy fitness function called a sampling fitness function utilizes sampling in order to reduce the amount of noise from fitness evaluations in noisy environments. In accordance with the Central Limit Theorem, the sampling fitness function reduces the amount of noise by taking the mean of multiple noisy fitness function evaluations for a given trial design. This Monte Carlo approach to fitness function evaluation

provides a more accurate estimate of the fitness of a given trial design because the design is exposed to a wide variety of conditions.

Recent research by *Miller* [1997] investigated the effects of noise on GA convergence rates and population sizing in order to determine the optimal sample size (i.e., the number of noisy fitness function evaluations) for GAs that employ sampling fitness functions. GA performance is a function of both the convergence rate of the GA and the number of generations the GA can run within a computational time constraint. The optimal sample size is the size at which the performance gains from faster convergence rates and decreased population sizes balances the performance loss due to larger sample sizes. One of the main points of guidance provided by the research described in *Miller* [1997] is given as follows: when the computational time required to perform a single fitness function evaluation for a given trial design dominates the amount of computational time required for the operations of selection, crossover, and mutation on one individual in one generation (e.g., as in this application, when a computationally-intensive simulation model is used to evaluate the objective function and constraint violations), the sample size (or the number of noisy fitness function evaluations for a given trial design) should be limited to a size approaching one.

Other research into sampling strategies for noisy GAs was performed by *Aizawa and Wah* [1994]. The authors investigated the effect of varying the duration of each generation and the number of noisy fitness function evaluations (sample size) for each trial design, given a fixed population size. Regarding GA performance, *Aizawa and Wah* [1994] assert that, in general, smaller sample sizes are better for early generations, but larger sizes are better for later generations. In their research, the authors developed methods to address what they term the "duration-scheduling" problem: all trial designs within a given generation are sampled equally,

but the sample size used to evaluate each design (and therefore the duration of each generation) is allowed to vary from one generation to the next.

2.7.3 Noisy Genetic Algorithm Operation

As mentioned in Section 2.6.3, multiple parameter sample sets are defined in this work, where each set consists of a single sample drawn randomly from the pool of generated conditional realizations and from each variable exposure model parameter distribution. As with Monte Carlo simulation, the sample sets for one trial design are formed independently from the sets for another trial design. Once *n* sample sets are formed for all trial designs within a generation, each trial design is simulated and evaluated in terms of the objective function and constraints. The overall fitness of a given trial design is then computed to be the average of the fitness evaluations for all *n* sample sets.

To determine the number of sample sets, *n*, the following approach for increasing the number of sample sets over the course of an optimization run, in the spirit of *Aizawa and Wah* [1994], was developed by trial-and-error. Although *Miller's* [1997] results showed that a sample size of 1 should be sufficient, initial trials indicated that this sample size did not produce designs of sufficient reliability. Instead, for the first four generations, *n* was set to 5 and was increased by five sample sets every four generations. Following generation twelve, the fittest four designs from the previous four generations (i.e., generations 9-12) were tested by simulating each with 500 sample sets. If any of the four designs were successful in meeting the specified reliability criterion, then *n* was not increased and the optimization process continued for four more generations before termination. Otherwise, *n* continued to be increased by five sample sets in the manner described above with a test of the four fittest designs from the previous four generations every four generations until successful termination or until a maximum number of generations

was reached. The implicit assumption is that *n* is increased during the optimization process to a number large enough to produce designs that are sufficiently reliable and that 16 generations (the minimum number possible under the current methodology) is sufficient to produce designs that are optimal or near-optimal. For the application shown in Section 3, the convergence results indicated this was a reasonable assumption. While this approach performed satisfactorily, further research is needed to determine the best sampling strategy and termination criterion. The continued application of noisy GA theory, developed through the research of *Miller* [1997], *Aizawa and Wah* [1994], and others, to complex optimization problems will provide a better understanding of how to formulate efficient noisy GAs.

3. Management Model Application

To demonstrate an implementation of the management model described in Section 2, a realistic case study was developed from published data. To the extent possible, the data were taken from the Borden site [*Graham and McLaughlin*, 1991; *Mackay et al.*, 1986]; deviations were made only when information was incomplete or nonexistent or when necessitated by the form or function of this management model. A plan-view schematic of the aquifer is shown in Figure 2, including the numerical grid, boundary conditions, and the location of the initial contaminant plume, the potential remediation well sites, and the monitoring wells. The parameter values chosen to represent this site within the numerical, analytical, and exposure and risk assessment models are given in Sections 3.1-3.3. Parameter values used in the objective function are given in Section 3.4. Section 3.5 describes parameters used in the noisy genetic algorithm.

3.1 Numerical Model Parameters

The unconfined, heterogeneous, isotropic aquifer shown in Figure 2 has dimensions of approximately 60 m by 20 m and was modeled using the grid of 16 by 8 elements shown. This grid was derived from a finer numerical grid of 128 by 64 elements, which was used to generate the conditional hydraulic conductivity realizations (see Appendix C). Following conditional simulation, the coarse grid was created by arithmetically averaging the hydraulic conductivity values using the averaging procedure illustrated in Figure 3, where each dot represents a hydraulic conductivity value for an element of the numerical contaminant transport simulation model grid. Some loss of accuracy inevitably results from the averaging procedure, but it was assumed that such losses were acceptable compared to the gains in computational speed-up.

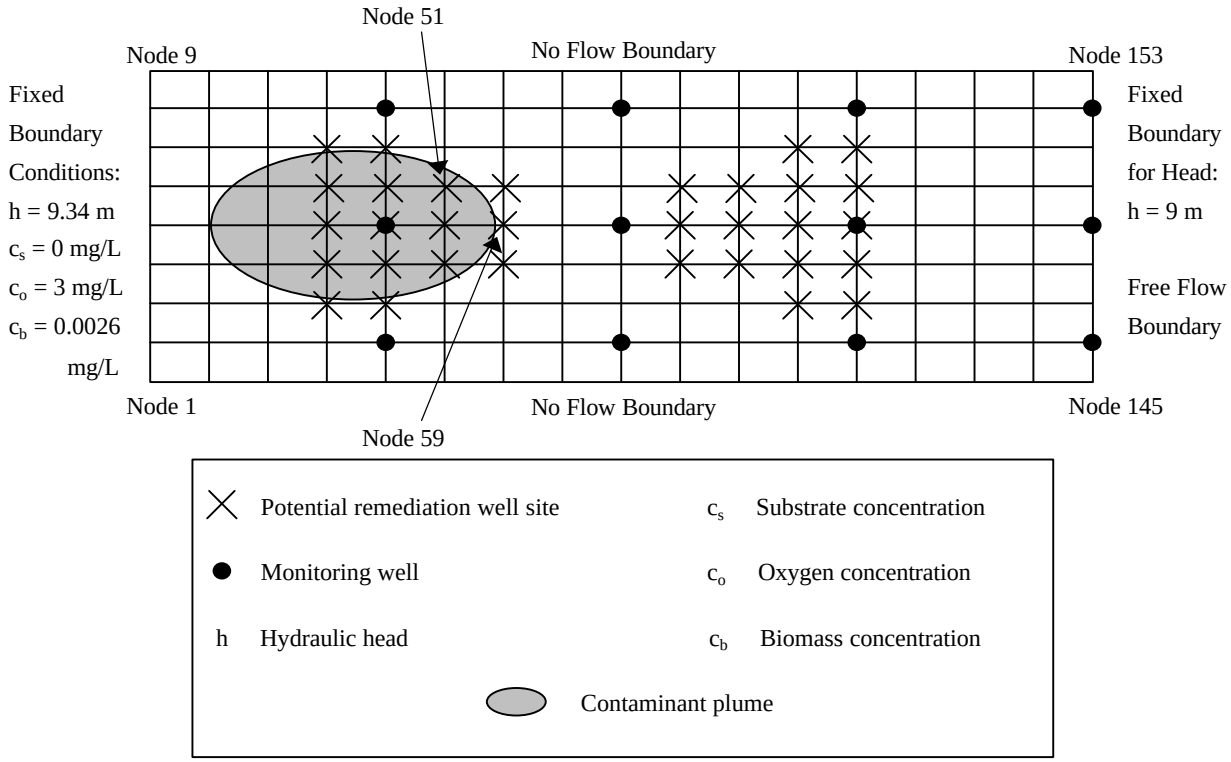

Figure 2. Plan-View Schematic of Aquifer

In order to quantify the effects of hydraulic conductivity averaging, Bio2D simulations with the coarse and fine grids were performed under the same conditions and without pumping. A single hydraulic conductivity realization that was assumed to be the actual distribution at the site was used for the two simulations. The percent deviations in steady-state hydraulic head [(averaged value – unaveraged value)/unaveraged value] at all nodes where a comparison was possible between the two simulations were compiled. The differences in hydraulic conductivity values for the two simulations are relatively small: the largest deviation was approximately 0.0263% and most deviations were negligible.

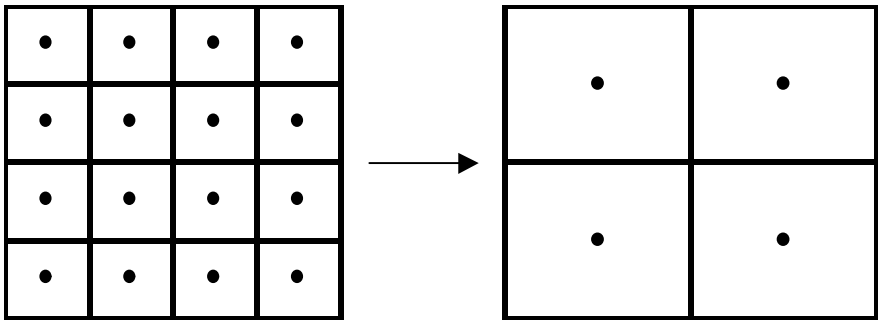

Figure 3. Method for the Averaging of Log-Hydraulic Conductivity Values

Other aquifer and operating parameters used in the numerical model are summarized in Table 1 (the variables shown refer to Equations (20)-(26) in Appendix A). For this application, the contaminant was assumed to be benzene, a common petroleum constituent for which RBCA procedures are well established. The hydraulic conductivity statistical parameters K_G, s_f^2, and l_K given in Table 1 were used to generate the conditional realizations (see Section 2.6.1 and Appendix C). For this application of the management model, there were 1,000 realizations in the pool of conditional realizations (see Appendix C). The mean hydraulic gradient, *J*, was used to establish the upgradient and downgradient hydraulic head boundary conditions. Although the Borden aquifer is unconfined, the formulation of Bio2D assumes a confined aquifer (see Section 2.3 and Appendix A), so the aquifer thickness, *b*, represents the magnitude of the undisturbed, unconfined hydraulic head field. The hydraulic head limits [$h_{min,l}$ and $h_{max,l}$ in Equation (8)] were set to values to ensure that disturbances to the hydraulic head field due to pumping were limited and that the confined aquifer assumption would remain reasonable (see Section 3.5 for these values and Appendix A for more discussion on the confined aquifer assumption).

The dispersivities (a_x and a_y) in Table 1 are within the range of values for sand given in *Freeze and Cherry* [1979]; for numerical accuracy, the chosen dispersivities and grid spacing

were checked to ensure a grid Peclet number less than 1. Likewise, the time step size and grid spacing shown in Table 1 were chosen to ensure a Courant number less than 1 for the groundwater velocities expected to occur with pumping. The concentration of oxygen in the injection water, $c_o^{'}$, is approximately the saturation value for air in water and is assumed to be achieved through aeration. The concentrations of substrate (contaminant) and biomass in the injection water, $c_s^{'}$ and $c_b^{'}$, were initially zero but were then set to the current concentration in the extracted water (adjusting for dilution if the total injection rate was greater than the total extraction rate) since all extracted water was re-injected. The substrate retardation factor (R_s) in Table 1 was calculated using the porosity, soil bulk density, and substrate partition coefficient shown [see Equation (24) of Appendix A], which were derived from values for sandy soils and benzene. Similarly, the biomass retardation factor (R_b) in Table 1 was calculated from the porosity, soil bulk density, and biomass partition coefficient shown [see Equation (25) of Appendix A].

Table 1. Bio2D Aquifer, Operating, and Biodegradation Parameters

Parameter	Value	Reference
Contaminant	Benzene	
Mean hydraulic conductivity (K_G), m/day	6.18	*Graham and McLaughlin* [1991]
Variance of lnK (s_f^2)	0.29	*Graham and McLaughlin* [1991]
lnK integral scale (l_K), m	2.8	*Graham and McLaughlin* [1991]
Mean hydraulic gradient (J)	0.0056	*Graham and McLaughlin* [1991]
Porosity (q)	0.33	*Graham and McLaughlin* [1991]
Aquifer thickness (b), m	9	*Mackay et al.* [1986]

Table 1. (continued)

Longitudinal dispersivity (a_x), m	3.78	
Transverse dispersivity (a_y), m	3.77	
Time step (Δt), days	0.5	
Grid spacing (Δx), m	3.8	
Grid spacing (Δy), m	2.5	
Oxygen concentration in injection water $\left(c_o^{'}\right)$, mg/L	10	
Soil bulk density (r_b), g/cm^3	2.00	*Lyman et al.* [1992]
Substrate partition coefficient (K_d), cm^3/g	0.062	*Lyman et al.* [1992]
Biomass partition coefficient (K_b), cm^3/g	15.26	*Dosa* [1994]
Substrate retardation factor (R_s)	1.38	
Biomass retardation factor (R_b)	62.96	
Substrate half-saturation coefficient (K_s), mg/L	3.99	*Chang et al.* [1993]
Oxygen half-saturation coefficient (K_o), mg/L	1	*Dosa* [1994]
Substrate inhibition coefficient (K_i), mg/L	10^{20}	
Maximum specific growth rate for biomass ($m_{\max}$), 1/day	4.95	*Chang et al.* [1993]
Yield coefficient (Y_c), mg/mg	1.13	*Chang et al.* [1993]
Ratio of oxygen to substrate used (F_{os})	3.077	
Endogenous respiration rate for bacteria (r_b), 1/day	0.05	*Dosa* [1994]
Background organic carbon concentration (c_c), mg/L	715	*Dosa* [1994]
Background organic carbon utilization rate (k_c), 1/day	10^{-5}	*Dosa* [1994]

The values for the Monod kinetic parameters (K_s, K_o, $m_{\max}$, and Y_c) shown in Table 1 are values found in the literature for benzene. The value of 10^{20} for the substrate inhibition coefficient assumes that inhibition effects are negligible. F_{os}, the ratio of oxygen to substrate used, was calculated from the stoichiometry for the complete aerobic biodegradation of benzene to carbon dioxide and water. The endogenous respiration rate for bacteria, the background organic carbon concentration and utilization rate, and the initial biomass concentration were chosen so that the biomass concentration would maintain or reach a steady-state in the absence of substrate (contaminant); these values assume a stable population of indigenous microorganisms supported by the background organic carbon.

The initial benzene concentration profile is depicted in Figure 4 and shows an initial peak concentration of approximately 133 mg/L. Initial oxygen and biomass concentrations were assumed to be 3 mg/L and 0.0026 mg/L, respectively, at all nodes. Also, benzene, oxygen, and biomass concentrations were assumed to be evenly distributed in the vertical direction, so that the use of the two-dimensional vertically averaged model was appropriate (see Appendix A).

The monitoring network shown in Figure 2 consists of four sets of monitoring wells with three wells to a set (M = 12); each set is at a specified x-coordinate location within the domain. At each point in time, t, during the remediation period, benzene concentrations were checked at the three downgradient boundary monitoring wells to ensure that the contaminant plume did not migrate beyond the monitored zone. Likewise, at the end of each monitoring period, assumed to be 90 days for this application, contaminant concentrations were also checked at each of the three non-boundary monitoring well sets. The analytical and exposure and risk assessment models were solved using the maximum concentrations observed among each monitoring well set, using analytical model inputs detailed in the next section, to produce estimates of $Risk_{t,k}^{TOTAL}$

for Equation (5). The simulation of a given remediation proceeded while the risk constraint, Equation (5), was satisfied for the boundary monitoring well set and continued until Equation (5) was satisfied at the end of a monitoring period for all four monitoring well sets. For this application, the maximum amount of time allowed to achieve successful risk-based remediation was set to five years. Also, to save computational effort the simulation was halted as soon as the risk constraint was violated at any time *t* at the boundary monitoring well set.

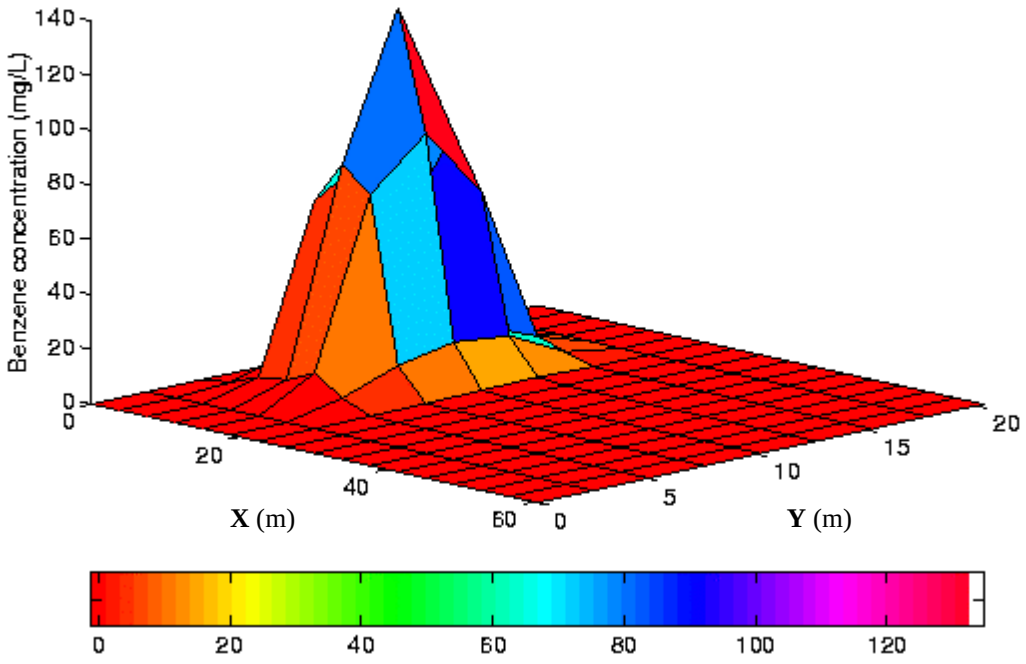

Figure 4. Initial Benzene Concentration Profile

3.2 Analytical Model Parameters

For each of the four monitoring well sets shown in Figure 2, the maximum concentration observed at time *t* provided the contaminant concentration input to the analytical model, C_t^{source}, in Equation (11). A single drinking water well, with water extracted for human consumption and use, was assumed to be located 50 m downgradient from the center of the right domain boundary shown in Figure 2. This well was assumed to be the sole exposure point for this application. The value of x_k in Equation (11), then, was the distance in the x-direction from each monitoring well set to the drinking water well, and S_{wid} was assumed to be 15.23 m, the extent in the y-direction of each monitoring well set. The average linear groundwater velocity in Equation (11), v, was estimated by taking the mean of the x-component velocities across the numerically modeled domain for a given hydraulic conductivity conditional realization without pumping. The first-order degradation constant [l_b in Equation (11)] for benzene was 0.007/day, from a study of the aerobic biodegradation of aromatic hydrocarbons at the Borden site [*Barker et al*., 1987]. Finally, the values for the dispersivities (a_x and a_y) for the numerical model shown in Table 1 were also used as inputs to the analytical model.

3.3 Exposure and Risk Assessment Model Parameters

Table 2 summarizes the single-point parameter values used in the exposure and risk assessment model [Equations (12)-(17) and (27)-(28)] which were derived from U.S. EPA's default exposure assumptions or chosen from ranges of values found in the literature. Table 3 gives the means and standard deviations obtained from the literature for those parameters in Equations (12)-(17) and (27)-(28) assumed to be variable. All of the variable parameters were assumed to follow lognormal distributions, although the normalized distribution parameters are

given in Table 3. Also, the last five parameters given in Table 2 and the last two parameters in Table 3 were used only to calculate the contaminant concentrations in air, $C_{t,k}^{a,shw}$ and $C_{t,k}^{a,nc}$ (see Appendix B).

The distribution parameter values given in Table 3 for exposure duration are the same for both the adult exposure duration, ED_{ad}, and the child exposure duration, ED_{ch} [used in Equations (15)-(17)]. The distribution for child exposure duration, though, was truncated at six years [for reference, see *Smith*, 1994]. Likewise, the values for adult air inhalation rate, $IR_{a,ad}$, and the child air inhalation rate, $IR_{a,ch}$, in Equations (16)-(17) are derived from the same air inhalation rate, IR_a, shown in Table 3, by multiplying IR_a by the body weight for an adult (BW_{ad}) or child (BW_{ch}), respectively. Finally, it should be noted that the adult body weight distribution, BW_{ad}, in Table 3 was derived by combining the distributions for both males and females using Monte Carlo simulation.

Table 2. Single-Point Exposure and Risk Assessment Model Parameters

Parameter	Value	Reference
Oral cancer potency slope (CPS_o), $(mg/day\text{-}kg)^{-1}$	0.029	*U.S. EPA* [1997]
Inhaled cancer potency slope (CPS_i), $(mg/day\text{-}kg)^{-1}$	0.029	*U.S. EPA* [1997]
Residential exposure frequency (*EF*), days/year	350	*U.S. EPA* [1991]
Child body weight (BW_{ch}), kg	15	*U.S. EPA* [1991]
Averaging time (*AT*), days	25,550	*U.S. EPA* [1991]
Time spent indoors (ID), days/day	0.67	*U.S. EPA* [1991]
Volatilization factor (*f*)	0.9	*Schaum et al.*[1994]
Bathroom volume (*BV*), m^3	6	*Schaum et al.*[1994]

Table 2. (continued)

House volume (*HV*), m^3	200	*Schaum et al.*[1994]
House exchange rate (*ER*), 1/day	10	*Schaum et al.*[1994]
Mixing coefficient (*MC*)	0.3	*Schaum et al.*[1994]

Table 3. Variable Exposure and Risk Assessment Model Parameters

Parameter	Mean	Standard Deviation	Reference
Adult drinking water ingestion rate ($IR_{w,ad}$), mL/day	7.02	0.49	*Roseberry and Burmaster* [1992]
Child drinking water ingestion rate ($IR_{w,ch}$), mL/day	6.43	0.50	*Roseberry and Burmaster* [1992]
Exposure duration (*ED*), years	1.98	0.95	*Israeli and Nelson* [1992]
Adult body weight (BW_{ad}), lb	5.04	0.13	*Brainard and Burmaster* [1992]
Air inhalation rate (IR_a), m^3/day-kg	-1.39	0.97	*ICRP* [1975]
Shower duration (*SD*), hours/day	-2.22	0.60	*James and Knuiman* [1987]
Time in bathroom after shower (BD), hours/day	-1.30	0.61	*McKone and Bogen* [1991]
Shower flow rate (SR_w), L/hour	6.12	0.33	*James and Knuiman* [1987]
Total household water use rate (HR_w), L/hour	3.62	0.36	*McKone and Bogen* [1991]

3.4 Objective Function Parameters

In Equation (2), the capital cost associated with the installation of each remediation well *i*, $c_i^{cap,rem}$, and the operating cost for remediation well *i*, $c_i^{op,rem}$, were derived from an analysis of well cost data performed at the Illinois State Water Survey [*Ackermann*, 1969a, 1969b]. The exact form of the equations used to estimate $c_i^{cap,rem}$ and $c_i^{op,rem}$ can be written as

$$c_i^{cap,rem} = a_0(d_{i,l})^{b0} + a_1(|u_{i,l}|)^{b1}(g_{i,l})^{b2} \quad (18)$$

$$c_i^{op,rem} = a_2(|u_{i,l}|)(g_{i,l}) \quad (19)$$

where $d_{i,l}$ is the depth of remediation well i at location l (m); $g_{i,l}$ is $\left(h_{i,l} - h_{i,l}^{und}\right)$ if $u_{i,l}$ is positive (injection) and is $\left(d_{i,l} - h_{i,l}\right)$ if $u_{i,l}$ is negative (extraction) (m); $h_{i,l}^{und}$ is the undisturbed hydraulic head for remediation well i at location l (m); and a_0, a_1, and a_2 and b_0, b_1, and b_2 are cost coefficients and cost exponents, respectively. The capital cost of remediation well installation, $c_i^{cap,rem}$, consists of two parts: a drilling cost and an equipment cost.

The values for the cost coefficients, a_0-a_2, and cost exponents, b_0-b_2, were taken from *Ackermann* [1969a, 1969b] and are summarized in Table 4. The values for the coefficients in Equations (3)-(4), except for $c^{muni,syst}$, were derived from environmental unit and assemblies cost manuals [*Rast*, 1997a, 1997b] and are also shown in Table 4. All cost coefficient values shown were inflated to a 1996 cost level using a 6% annual inflation rate. Also, future operating and other time-dependent costs in Equations (2)-(4) were discounted to present value using an assumed annual discount rate of 10%. An electricity cost of \$0.10/kWh was assumed when determining the values for a_0 and $c^{op,syst}$. The capital costs associated with monitoring well installation, $c^{capsite,mon}$ and $c_j^{capwell,mon}$, do not influence decisions regarding the optimal remedial design since they are affected only by the number of monitoring wells, chosen a priori by the user, and are in no way influenced by decision variable values. These costs were included only to improve the accuracy of the total remedial cost.

Table 4. Objective Function Cost Coefficients and Exponents

Coefficient/Exponent	Value	Reference
a_0	6923	*Ackermann* [1969a]
a_1	41.75	*Ackermann* [1969a]
a_2	1.09×10^{-3}	*Ackermann* [1969b]
b_0	0.299	*Ackermann* [1969a]
b_1	0.453	*Ackermann* [1969a]
b_2	0.642	*Ackermann* [1969a]
$c^{capsite,mon}$, \$	1,893.92	*Rast* [1997a, 1997b]
$c_j^{capwell,mon}$, \$	3,728.46	*Rast* [1997a, 1997b]
$c^{samp,mon}$, \$	392.95	*Rast* [1997a, 1997b]
$c_j^{ana,mon}$, \$	116.66	*Rast* [1997a, 1997b]
$c^{cap,syst}$, \$	411.66	*Rast* [1997a, 1997b]
$c^{muni,syst}$, \$/m^3	0.47	
$c^{op,syst}$, \$/day	151.67	*Rast* [1997a, 1997b]
$c^{main,syst}$, \$	1,034.94	*Rast* [1997a, 1997b]

3.5 Noisy Genetic Algorithm Parameters

Figure 2 shows the potential locations for remediation wells, which were grouped into two sets of sixteen locations, one upgradient and the other downgradient, for a total of thirty-two potential well locations. For this application, the total number of remediation wells that could be installed for any given remedial design was set to N = 2. A decision variable for pumping scheme was implemented in the genetic algorithm using an eight-digit binary string representing three possible schemes: (1) an injection well located within the upgradient well set and an

extraction well located within the downgradient well set, (2) an extraction well within the upgradient set and an injection well within the downgradient set, or (3) an injection well within the upgradient set and another within the downgradient set. The installation of two extraction wells was not included as a possibility because any extracted water must be re-injected (see Section 2.2), but the absence of a remedial system (no wells installed) was a possible decision. The exact location of a remediation well within a particular well set, *l*, was determined by another decision variable, represented by a four-digit binary string that designates one of the sixteen potential locations.

The remediation well pumping rates, $u_{i,l}$, selected by the genetic algorithm were constrained to lie within the range of u_{min} = 0 to u_{max} = 64 m^3/day and were represented by six-digit binary strings. If, however, an extraction rate was selected that was greater than that for injection for either of the first two scenarios described above, the extraction rate was set to be equal to that for injection. The value for u_{max} allowed a significant range of remedial designs to be investigated, while also minimizing the number of potential head constraint violations. The values of $h_{min,l}$ and $h_{max,l}$ were set to ensure that changes in the hydraulic head at well locations *l* due to pumping were no greater than 10% of the undisturbed head value. Finally, the well installation decision variables, x_i, were each represented by a single bit. Well installation was explicitly included as a decision variable (rather than an implicit function of the pumping rates selected) because the GA performance was substantially better under this approach.

As described in Section 2.7.3, the "fittest" four designs from the previous four generations were tested periodically by simulating each with 500 sample sets in order to determine whether *n*, the number of sample sets, was large enough. For this application, the reliability criterion specified that, for this test, there can be no head constraint violations and at

least 90% of the 500 sample sets must have no risk constraint violations. Penalty functions, as described in Section 2.7.1, were used to enforce the risk constraint, Equation (5), and the head constraint, Equation (8). The penalty coefficients for the risk and head constraints were set to values of 3.33×10^{9} and 10^{6}, respectively. Also, a population size of 150 was used, and the maximum number of generations allowed was 24. Finally, the probabilities of crossover (*Pc*) and mutation (*Pm*) were set to 0.9 and 0.0067, respectively.

4. Results and Discussion

The management model was solved twice with the target risk (TR) specified at levels of 10^{-5} and 10^{-6}, respectively, for the two solutions. For both solutions, the noisy GA proceeded through 16 generations prior to termination. As mentioned in Section 2.7.3, the convergence results indicated that the assumption of a minimum number of 16 generations was sufficient to produce designs that are optimal or near optimal. During the optimization process, there were consistent increases in the population fitness mean and decreases in the population fitness standard deviation for both solutions as the GA progressed from one generation to the next. Furthermore, by the final generations the fitness of the best design differed from the population fitness mean by approximately $5,000 and $10,000 for the two solutions. Most of the difference between the population best and mean fitness in the final generations is attributable to the aberrant and extremely "unfit" designs that are sometimes produced, even in later generations, when using such a large population size.

Following optimization, each design determined to be optimal for each TR level was evaluated by simulating each with 1,000 sample sets, formed randomly from the pool of conditional realizations and from the exposure parameter distributions, in order to assess the reliability of each design. The management model solutions for the two TR levels as well as the results of the reliability evaluation are presented in Table 5. For both TR levels, the optimal design found by the management model was to locate a single injection well within the upgradient well set, injecting at the rates given in Table 5. Results of the reliability evaluation indicate that both designs exhibit a high degree of reliability: the design for the TR level of 10^{-5} failed to meet the target risk only once, whereas the design for the TR level of 10^{-6} had no risk constraint violations (neither design violated any head constraints). These high levels of reliability were achieved by the noisy GA with a maximum of only 15 sample sets selected from

the hydraulic conductivity conditional realizations and exposure parameter distributions. These results indicate the efficiency of the noisy GA in finding reliable designs with only a coarse level of information on uncertainty.

Note in Table 5 that the mean risk level achieved by each design is more than an order of magnitude less (more protective of human health) than the TR level specified for each management model run due to the desire for high design reliability. These results demonstrate that uncertainty regarding the spatial distribution of hydraulic conductivity and exposure parameter variability significantly influence the assessment of risk. Therefore, understanding the uncertainty associated with risk assessment can be just as important as the decision regarding the appropriate target risk level. It should also be noted that of the three components that comprise total risk in this management model, the risk due to the inhalation of volatiles resulting from non-consumptive uses of water in the home ($Risk^{nc}$) is the dominant contributor to the total risk assessed.

Table 5. Risk-based In Situ Bioremediation Design Variables and Reliability Assessment Results

Design Variable/Result	Target Risk (TR)	
	10^{-5}	10^{-6}
Injection Pumping Rate(s), m^3/day	11.175	31.492
Injection Well Location(s), node #	51	59
Extraction Pumping Rate, m^3/day	NA	NA
Extraction Well Location, node #	NA	NA
Number of Generations Before Termination of GA	16	16
Design Reliability, %	99.9	100

Table 5. (continued)

Total Risk Level ($Risk^{TOT}$) (maximum value among all four monitoring well sets)	Mean	9.424×10^{-7}	8.385×10^{-8}
	Standard Deviation	1.592×10^{-6}	1.460×10^{-7}
Risk Level due to Ingestion ($Risk^{w}$)		7.938×10^{-8}	6.820×10^{-9}
Risk Level due to Showering ($Risk^{shw}$)		1.160×10^{-8}	1.018×10^{-9}
Risk Level due to Other Non-Consumptive Uses of Water ($Risk^{nc}$)		8.514×10^{-7}	7.601×10^{-8}
Remediation Period, days	Mean	92.70	91.53
	Standard Deviation	18.70	11.64
Total Present Value Cost of Remediation (C_{TOT}), $	Mean	65,405	66,302
	Standard Deviation	820	612
Present Value of Remediation Well Costs (C_{REM}), $		14,225	14,321
Present Value of Monitoring Costs (C_{MON}), $		48,436	48,414
Present Value of Additional Remedial System Costs (C_{SYST}), $		2,744	3,567

Another significant result shown in Table 5 is that the mean total remediation costs (C_{TOT}) for the two different solutions differ by less than $1,000. In order to meet a target risk level that is an order of magnitude more protective, the injection rate only needed to be increased by approximately 20 m^3/day, and this added operating cost is relatively small compared to the total remediation cost. It can also be seen in Table 5 that the standard deviations of the total remediation costs ($820 and $612 for TR levels 10^{-5} and 10^{-6}, respectively) approach in magnitude the difference between the mean total remediation costs (C_{TOT}) of the two solutions (~$900). This result further illustrates the importance of understanding the uncertainty associated with the risk assessment process, since the total remediation cost can be affected by this uncertainty as much as by the target risk level that is chosen.

Table 5 also lists the individual cost components that comprise the total remediation cost, C_{TOT}. Although the costs associated with monitoring, C_{MON}, are a majority of the total remediation cost, most of C_{MON} consists of capital costs associated with installing the monitoring wells, the number of which is chosen a priori. Because the noisy GA determined fairly early in the optimization process that only a single injection well was necessary for successful remediation, the main evaluation made by the noisy GA during the entire optimization process was not how many remediation wells to install, but how to trade off the operating costs that comprise C_{REM}, C_{MON}, and C_{SYST}. For example, a lower injection rate decreases remediation well operating costs but also leads to longer remediation periods, thereby increasing the time-dependent cost associated with C_{MON}. C_{SYST} consists of costs that are a function of both the injection rate and time; therefore, the effect of a decreased injection rate on C_{SYST} can not be established in a general sense. However, the time-dependent costs associated with C_{MON} and C_{SYST} were significant for this application because the noisy GA found the minimum injection rates for the two TR levels that achieved successful remediation by the end of the first monitoring period (90 days). As shown in Table 5, the mean remediation periods for the two TR levels were 92.70 days and 91.53 days, respectively. These results indicate the potential importance of considering monitoring and other remediation system operation and maintenance costs in RBCA design, even though such costs are not directly affected by the pumping rates selected. Hence, allowing a variable remediation period so that tradeoffs among these costs can be considered is clearly an important model feature for this application.

5. Conclusions and Further Research

A methodology for developing cost-effective, reliable in situ bioremediation designs within a risk-based framework was presented and applied to a realistic case study. The method couples a noisy genetic algorithm with numerical and analytical fate and transport models for predicting concentrations in the source zone and at the exposure well, respectively. An exposure and risk assessment model then translates the predicted concentrations into estimates of human health risk, allowing risk-based criteria to be considered in the corrective action design. A major contribution of this work was the implementation of a noisy GA to incorporate the uncertainty regarding the spatial distribution of hydraulic conductivity and the variability of exposure parameters within the search for optimal remedial designs. An evaluation of the optimal designs demonstrated that the noisy GA was capable of generating designs that exhibited a high degree of reliability from a relatively small number of sample sets, a significant advantage for computationally-intensive groundwater management models.

While this research has demonstrated the advantages of using noisy GAs to incorporate uncertainty, further work remains to understand how noisy GAs can be implemented most effectively within groundwater management models. Additional research is needed to understand how sampling strategies and termination criteria affect GA efficiency, the values of the decision variables for the optimal design, and the reliability of the optimal design. Research is ongoing to investigate these issues and to develop more computationally efficient methods for applying the model to large-scale field sites.

The results of the management model application demonstrated that modeling uncertainty can significantly influence the assessment of risk. An understanding of the degree of uncertainty in the risk assessment process is important when making tradeoffs between remediation cost and the appropriate target risk level. For the application shown, meeting a target risk level that is an

order of magnitude more protective did not significantly increase the mean total remediation cost because the additional pumping was inexpensive relative to the other remediation costs. Also, it was shown that the time-dependent costs associated with monitoring and the remedial system were significant for this application of the management model, illustrating the potential importance of variable remediation periods to allow tradeoffs among these costs and pumping rates to be considered. Further research is needed to investigate whether these findings will also apply to field-scale applications.

Although this work assumed that in situ bioremediation would be the remediation strategy, the model could easily be modified in future research to allow evaluation of any combination of pump-and-treat, containment, natural attenuation, and in situ bioremediation. A significant contribution of this work toward allowing tradeoffs among these technologies to be considered is the inclusion of a variable remediation period for meeting the risk constraints. To add pump-and-treat and containment to the model, only a cost component for aboveground treatment of extracted water would need to be added. To add natural attenuation, other electron acceptors such as nitrate, iron, and sulfate, which can be critical to the performance of natural attenuation, should be added to the model. Moreover, the computational efficiency of the model would need to be improved before long-term natural attenuation can be considered. Research is ongoing in these areas. Other enhancements to this management model could include the use of statistical distributions to characterize other variable exposure parameters that were assigned single-point values in this management. This would provide a greater understanding of the uncertainty inherent in the risk assessment process.

Although future work is needed, this research has provided a general methodology for developing cost-effective, reliable groundwater remediation designs within a risk-based

framework. Risk-based corrective action has become a widespread approach for sites contaminated from petroleum releases; ASTM is currently developing risk-based standards that can be applied at any chemical release site. As such guidelines take hold, tools such as the one described in this work will be valuable for groundwater planning and management under a RBCA approach.

Appendices

Appendix A. Bio2D

The numerical simulation model is a two-dimensional, depth-averaged finite-element flow, transport, and biodegradation model called Bio2D [*Taylor*, 1993]. The model predicts hydraulic heads and substrate (contaminant), oxygen, and microbial biomass concentrations in a confined aquifer that would result from pumping rates and well locations selected by the optimization component of the management model. The four governing equations that characterize this model are given in Equations (20)-(23):

$$\nabla \cdot (b\mathbf{K}_h \nabla h) + \sum_{i=1}^{N} u_{i,l} d_l = 0 \tag{20}$$

$$R_s bq \frac{\partial c_s}{\partial t} = \nabla \cdot (b\mathbf{D} \cdot \nabla c_s) - bv \cdot \nabla c_s - bqc_b R_b r_s + \sum_{i=1}^{N} u_{i,l} (c_s^{'} - c_s) d_l \tag{21}$$

$$bq \frac{\partial c_o}{\partial t} = \nabla \cdot (b\mathbf{D} \cdot \nabla c_o) - bv \cdot \nabla c_o - bqc_b R_b F_{os} r_s + \sum_{i=1}^{N} u_{i,l} (c_o^{'} - c_o) d_l \tag{22}$$

$$R_b bq \frac{\partial c_b}{\partial t} = \nabla \cdot (b\mathbf{D} \cdot \nabla c_b) - bv \cdot \nabla c_b + bq \left[c_b R_b (Y_c r_s - r_b) + c_c Y_c k_c \right] +$$

$$\sum_{i=1}^{N} u_{i,l} (c_b^{'} - c_b) d_l \tag{23}$$

Equation (20) describes steady-state flow and is solved once for a set of pumping rates, $u_{i,l}$ (m^3/day). In Equation (20), b is the aquifer thickness (m), $\mathbf{K}_h$ is the hydraulic conductivity tensor (m/day), h is the hydraulic head (m), N is the number of remediation wells, and d is the Dirac delta function at remediation well location l. Equations (21)-(23) describe time-dependent changes in substrate (contaminant), oxygen, and microbial biomass concentrations due to advection, dispersion, adsorption and biodegradation. In Equations (21)-(23), c_s, c_o, and c_b are the concentrations of substrate (contaminant), oxygen, and biomass, respectively, in the aqueous phase (mg/L); R_s and R_b are the substrate and biomass retardation factors, respectively; q is the porosity; **D** is the dispersion tensor (m^2/day); v is the Darcy velocity (m/day), which is computed using Darcy's law [for reference see, for example, *Freeze and Cherry*, 1979]; $c_s^{'}$, $c_o^{'}$, and $c_b^{'}$ are the concentrations of substrate (contaminant), oxygen, and biomass, respectively, in the injection water (mg/L); F_{os} is the ratio of oxygen to substrate used; Y_c is the yield coefficient (mg/mg); r_b is the endogenous respiration rate for bacteria (1/day); c_c is the background organic carbon concentration (mg/L); and k_c is the background organic carbon utilization rate (1/day).

The substrate and biomass retardation factors in Equations (21)-(23), R_s and R_b, are calculated by

$$R_s = \left(1 + \frac{r_b K_d}{q}\right) \tag{24}$$

$$R_b = \left(1 + \frac{(1 - q)}{q} r_b K_b\right) \tag{25}$$

where r_b is the soil bulk density (g/cm^3), K_d is the substrate partition coefficient (cm^3/g), and K_b is the biomass partition coefficient (cm^3/g). The rate of substrate (contaminant) biodegradation, r_s in Equations (21)-(23), is modeled using the Haldane variant of the Monod equation,

$$r_s = \frac{m_{\max}}{Y_c}\left(\frac{c_s}{K_s + c_s + \frac{(c_s)^2}{K_i}}\right)\left(\frac{c_o}{K_o + c_o}\right) \tag{26}$$

where K_s and K_o are the substrate and oxygen half-saturation coefficients, respectively (mg/L); K_i is the substrate inhibition coefficient (mg/L); and $m_{\max}$ is the maximum specific growth rate for biomass (1/day). The Haldane variant of the Monod equation given in Equation (26) allows inhibition of microbial growth to occur at high substrate (contaminant) concentrations and is discussed in more detail in *Gaudy and Gaudy* [1980].

Bio2D uses a Galerkin finite-element method in space and a variably weighted, backward-difference approximation in time to solve Equations (20)-(23). The finite-element formulation allows for spatial variation in hydraulic conductivity, aquifer thickness, dispersivities, and porosity by specifying different values for each element of the mesh. This flexibility is essential to the management model described here because simulations are performed using multiple heterogeneous hydraulic conductivity realizations.

Several simplifying assumptions are made within the formulation of Bio2D, which are mentioned briefly here and discussed in greater detail in *Minsker and Shoemaker* [1998]. First, the flow equation, (20), is in a steady-state form. It is assumed that the hydraulic heads adjust quickly to pumping relative to the length of the remediation period. For any given aquifer, the

transient form of this equation can be used if this assumption is not valid. A further assumption is the use of a confined aquifer formulation, Equation (20), to approximate the hydraulic behavior of an unconfined aquifer. Using a confined aquifer formulation for an approximation is often justified as long as disturbances to the hydraulic head field due to pumping are relatively small [see *Atwood and Gorelick*, 1985]. As above, only modifications to Equation (20) are necessary if this assumption is not valid. Another simplifying assumption is the modeling of flow and transport in two dimensions by depth averaging in the vertical direction. Depth averaging is justified for sites that are relatively homogeneous in the vertical direction. At the Borden site, hydrogeologic variability is relatively small [*Graham and McLaughlin*, 1991]. In *Graham and McLaughlin* [1991], the authors were able to use a two-dimensional, vertically averaged transport model to successfully predict the evolution of the Borden chloride plume. A final simplification in Bio2D is the fact that the effects of bacterial clogging on aquifer permeability and porosity are not included. If these effects are considered to be significant, biomass concentration constraints can be added to the management model to ensure that clogging does not occur.

An in-depth description and verification of Bio2D is given in *Dosa* [1994], including additional Bio2D assumptions such as: (1) ionic and molecular diffusion are negligible relative to hydrodynamic dispersion, (2) oxygen and substrate (contaminant) are the only rate-limiting factors to microbial growth; nitrogen, phosphorus and other trace organics are sufficiently available, (3) the complex subsurface bacterial population can be represented by a single facultative, heterotrophic bacterial type; in the absence of hydrocarbon, they maintain a constant background concentration by consuming the background carbon found in the aquifer, (4) the amount of background carbon, c_c, remains constant in the aquifer, (5) background organic carbon

utilization is given by a first-order rate constant, and (6) oxygen is recharged at a rate necessary to offset that consumed in the aerobic degradation of background carbon (thus no explicit term is included in Equation (22) for oxygen use during degradation of background carbon).

Appendix B. Calculation of Contaminant Concentrations in Air

The following presents and describes the equations used to calculate contaminant concentrations in shower/bathroom air ($C_{t,k}^{a,shw}$) and household air ($C_{t,k}^{a,nc}$), as needed for Equations (16) and (17) in the exposure model, given the contaminant concentration in groundwater extracted for human consumption and use ($C_{t,k}^{w}$). Equation components and assumptions are discussed but reference should be made to *Schaum et al.* [1994], the source for this inhalation exposure methodology, for a complete discussion.

Showering Exposure

This section develops a simple model for calculating contaminant concentrations in bathroom air. The model treats the bathroom as a single, closed compartment and produces an air concentration that is averaged over the time spent both in the shower and in the bathroom following the shower. In addition, this model is based on the assumptions that the contaminant volatilizes at a constant rate and that it instantaneously and uniformly mixes with the bathroom air. With these specifications and assumptions, the exposure concentration during showering and the time spent in the bathroom afterwards is provided by the following equation:

$$C_{t,k}^{a,shw} = \frac{C_{t,k}^{w} \cdot SR_w \cdot f \cdot SD}{BV} \cdot \left(\frac{SD/2 + BD}{SD + BD} \right) \tag{27}$$

where SR_w is the shower flow rate (L/hour), f is the volatilization factor (or transfer efficiency), SD is the shower duration (hours), BD is the time spent in the bathroom after shower (hours), and BV is the bathroom volume (m^3). Note that the units for SD and BD differ from those given in

Section 2.5. In Section 2.5, the two parameters represent rates, whereas here they are simply durations. The values for these parameters, though, remain the same. The assumptions of constant rate of volatilization and no ventilation (closed compartment) contribute to the conservative nature of this model. Other assumptions include: (1) the initial contaminant concentration in air is assumed to be zero; i.e., no contaminant concentration remains from previous showers by others or from other household uses of water and (2) risk factors (cancer potency slope or *CPS*), though derived from continuous exposure experiments, are appropriate for such intermittent exposure.

Additional Non-Consumptive Use Exposure

Schaum et al. [1994] provide an additional model for indoor residential exposure to volatiles which doesn't account for the higher, short-term exposures that may occur due to showering but which includes more continuous exposure by accounting for most non-consumptive uses of water such as dish washing, clothes washing and cooking. The objective of this model is to produce an indoor air concentration that is the spatial average over the house throughout the day. This concentration is estimated by the following equation:

$$C_{t,k}^{a,nc} = \frac{C_{t,k}^{w} \cdot HR_w \cdot f}{HV \cdot ER \cdot MC} \tag{28}$$

where HR_w is the total household water use rate (L/day), *HV* is the house volume (m^3), *ER* is the house exchange rate (air changes) (1/day), and *MC* is the mixing coefficient. This is also a fairly simple model, but unlike the showering model, this one includes an exchange rate term (open system) and a mixing coefficient, since uniform mixing is unlikely to occur. One notable

assumption is that the same volatilization fraction used for showering exposure can be used here; i.e., the volatilization fraction obtained for showering is the average of all water uses throughout the entire house.

Appendix C. Stochastic Approach and Conditional Simulation

Stochastic Approach

This section details the stochastic approach to groundwater flow and contaminant transport modeling, the theory of conditional simulation, and the methodology used to apply the theory to this research. As discussed earlier, hydrogeologic properties, such as hydraulic conductivity, vary through space, in some cases by an order of magnitude or more over short distances. Because sufficient sampling to determine such parameters is infeasible, it is impossible to fully characterize the spatial distribution. Therefore, any estimates of the hydraulic conductivity distribution inevitably will be uncertain, and this uncertainty will be reflected in groundwater simulation model predictions of contaminant concentrations.

To overcome the inability to fully characterize aquifer properties due to heterogeneity, a stochastic approach is implemented. Under this approach, the heterogeneous hydraulic conductivity is viewed as a random field described by a probability density function (Pdf). The actual hydraulic conductivity distribution of a given site is considered to be a single realization of the random hydraulic conductivity field. The first step is to characterize the statistical structure of the random hydraulic conductivity field. A complete probabilistic description of the random field would require knowledge of the joint probability density function for hydraulic conductivity at all points throughout the aquifer. Such a description, though, is not feasible because we do not have sufficient data to determine the joint distributions; i.e. the hydraulic conductivity distribution for any given aquifer represents only a single realization of the random field.

The assumptions of stationarity and ergodicity allow the random field to be characterized using data from a single realization. "Stationarity assumes that any statistical property of the

medium (mean, variance, covariance, a higher-order moment) is stationary in space, i.e., does not vary with a translation. It will be the same at any point of the medium" [*Marsily,* 1986]. Here, only the assumption of weak stationarity, stationarity of the first two moments, is made. Therefore, the mean (first moment) of the random field is constant through space, and the covariance (second moment) at two points in space depends only on the separation distance. The joint probability density function can be fully characterized by the first two moments when the Pdf is normal, or Gaussian. "Ergodicity implies that the unique realization available behaves in space with the same Pdf as the ensemble of possible realizations. In other words, by observing the variation in space of the property, it is possible to determine the Pdf of the random function for all realizations" [*Marsily*, 1986]. The ergodic hypothesis is valid under certain conditions that are not described in detail here.

The hydraulic conductivity in this research was assumed to follow a lognormal distribution. In addition, it was assumed that the correlation of the random hydraulic conductivity field is represented by an isotropic, exponential covariance function. Both of these assumptions are supported by the findings of *Hoeksema and Kitanidis* [1985]. Given the assumptions outlined in this and the above paragraph, the defining statistical functions for the random hydraulic conductivity field are as follows:

$$Z_i = \ln K_i \tag{29}$$

$$E[Z_i] = m_Z \tag{30}$$

$$Cov[Z_i, Z_j] = s_Z^2 \exp\{-|\Delta x_{ij}|/l_Z\} \tag{31}$$

where Z_i is the natural logarithm of the hydraulic conductivity, K_i, at point x_i, $E[\ ]$ is the expected value, $Cov[\ ]$ is the covariance, m_Z is the log-hydraulic conductivity mean, s_Z^2 is the log-hydraulic conductivity variance, l_Z is the log-hydraulic conductivity correlation length or integral scale, and $|\Delta x_{ij}|$ is the separation distance of Z_i and Z_j. Therefore, the random log-hydraulic conductivity field is statistically characterized by the parameters m_Z, s_Z^2, and l_Z. If the statistical parameters of the log-hydraulic conductivity field are known, or can be estimated, then probabilistic investigations of groundwater flow and contaminant transport can be performed.

Conditional Simulation Theory

Estimation procedures like kriging are often used to predict the spatial distribution of hydraulic conductivity given a set of sampling data and estimates of the statistical parameters described in the previous section. "In a nutshell, the kriging algorithm provides a minimum error-variance estimate of any unsampled value" [*Deutsch and Journel*, 1992]. Such a procedure, though, produces estimated maps of hydraulic conductivity that are "smoother" (lower variance and larger correlation length) than the actual site distribution, as indicated by the sampling data. The information produced by the kriging procedure is valuable in terms of showing global trends, but it is desirable to use hydraulic conductivity maps that exhibit more of the natural variability when modeling groundwater flow and contaminant transport.

Conditional simulation is a stochastic simulation process that allows for the generation of multiple realizations of the random hydraulic conductivity field that (1) are statistically similar to the actual site (or realization considered to be the actual site) and (2) respect known information

about the actual site distribution gathered through sampling (hence, they are "conditioned" by the known values at specific points). Conditional simulation was originally used to correct for the "smoothing" effect that results from kriging: "Typical conditional simulation algorithms trade the estimation variance minimization for the reproduction of a variogram/covariance seen as a model of spatial variability" [*Deutsch and Journel*, 1992]. In addition to re-gaining the variance that was lost through kriging, multiple conditional realizations provide a means to investigate the uncertainty of the spatial distribution of hydraulic conductivity.

The following description of conditional simulation theory is based on *Journel and Huijbregts* [1978]. Consider a vector Z of ln$K$ values for an actual site distribution and a vector Z_{KR} of ln$K$ values that are estimates from kriging. Z is unknown in reality and therefore the estimation error, $(Z - Z_{KR})$, is also unknown. Consider, though, the following decomposition for Z:

$$Z = Z_{KR} + (Z - Z_{KR}) \tag{32}$$

Similarly, consider a vector of ln$K$ values from an unconditional realization, Z_{UR}, that have the same statistical properties as Z. This distribution can also be sampled at the same locations used to obtain Z_{KR} and kriged to produce a vector of ln$K$ estimates, $Z_{UR,KR}$. As above, the same decomposition can be expressed

$$Z_{UR} = Z_{UR,KR} + (Z_{UR} - Z_{UR,KR}) \tag{33}$$

but in this case, all values are known. The idea of conditional simulation is to substitute $(Z_{UR} - Z_{UR,KR})$ for $(Z - Z_{KR})$ so that

$$Z_{CR} = Z_{KR} + (Z_{UR} - Z_{UR,KR}) \tag{34}$$

It can be shown that Z and Z_{CR} have the same statistical properties [see *Journel and Huijbregts*, 1978]. As the number of samples is increased, the conditional realizations, as can be expected, become more like the actual field. Even when the number of samples is increased, the field statistics are still preserved. As the number of samples is increased, Z_{KR} becomes more like Z, thereby increasing the variance of Z_{CR}. At the same time though, $Z_{UR,KR}$ more closely approximates Z_{UR}, thereby decreasing the perturbation, $(Z_{UR} - Z_{UR,KR})$, that is added to Z_{KR}.

Conditional Simulation Methodology

The conditional simulation procedure generated nodal log-hydraulic conductivity values, but the numerical simulation model requires elemental values. Hence, slightly different grids were used for the two procedures, as shown in Figure 5. The solid lines define the boundaries of the domain for the numerical simulation model and delineate the elements within this domain. Likewise, the dotted lines define the boundaries of the domain for the conditional simulation procedures and delineate the elements within this domain. The nodes on the dotted grid are at the center of each element of the numerical simulation model grid and designate the spatial locations at which log-hydraulic conductivity values are generated, sampled, and estimated. The value of log-hydraulic conductivity at each node was then assigned to its corresponding element of the numerical simulation model grid.

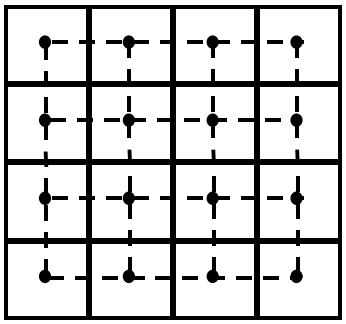

Figure 5. Numerical Simulation Model and Conditional Simulation Grids

The first step of the conditional simulation procedure is to generate the log-hydraulic conductivity realizations. For the application shown in Section 3, one realization, assumed to be the actual distribution, and 1,000 additional unconditional realizations, which are used to create the conditional realizations, were generated. A random field generator [*Gutjahr et al.*, 1995] was used to create realizations that are normally distributed with a mean of zero and a given variance and correlation length. The realizations have 128 nodes in the x-direction and 64 nodes in the y-direction and actual dimensions of 60m by 20m. For the correlation structure, an isotropic, exponential covariance function was employed.

The next step is to sample from the actual and 1,000 additional realizations. For this research, samples were chosen from each of the 1,001 realizations at 100 randomly selected nodes. The sample sets for each realization generally reflect the mean and variance specified for the random log-hydraulic conductivity field. Experimental semivariograms were then calculated for each of the sample sets. A FORTRAN program provided in GSLIB [*Deutsch and Journel*, 1992] was used to compute experimental semivariograms in the east-west direction. Although the semivariograms are in the east-west direction, the program was actually implemented in a

manner to compute “omni-directional” experimental semivariograms (i.e., data in the north-south direction is used) because it is assumed that there is no field anisotropy.

Model semivariograms were then fit to each of the 1,001 experimental semivariograms. An exponential model,

$$\gamma(h) = c^*[1-\exp(-h/a)] \tag{35}$$

where h is the separation distance, c is the sill, and a is the range, was used to fit the experimental semivariograms by minimizing the sum of squared errors between the experimental and model semivariograms, thereby providing values for c and a. For each of the experimental semivariograms, the model was fit to the experimental semivariogram only up to a separation distance of x = 30m.

Once the model parameters have been determined for the sample sets, the next step is kriging. Using the 1,001 sample sets and the model semivariogram parameters for each, an ordinary kriging FORTRAN program from GSLIB [*Deutsch and Journel*, 1992] with an isotropic, exponential covariance function was implemented to estimate ln$K$ at each node in the grid. The means for these estimated maps were similar to those for the generated realizations, but as expected, the variances were significantly lower. As mentioned, kriging generally produces estimated maps that are “smoother”, i.e. lower variance and longer correlation length.

The final step is to produce the conditional realizations. Using the expression described earlier for Z_{CR}, 1,000 conditional realizations were produced using the kriged map from the actual samples, the 1,000 unconditional realizations, and the 1,000 kriged maps from the samples of the 1,000 unconditional realizations. As predicted, the statistics of the conditional realizations

were similar to those of the random log-hydraulic conductivity field, in terms of both mean and variance.

Bibliography

Ackermann, W. C., Cost of Wells and Pumps, *Ground Water*, 7(1), 35-37, 1969a.

Ackermann, W. C., Cost of Pumping Water, Ground Water, 7(1), 38-39, 1969b.

Aizawa, Akiko N., and Benjamin W. Wah, Scheduling of Genetic Algorithms in a Noisy Environment, *Evolutionary Computation*, 2(2), 97-122, The Massachusetts Institute of Technology, 1994.

Andricevic, Roko, and Vladimir Cvetkovic, Evaluation of risk from contaminants migrating by groundwater, *Water Resources Research*, 32(3), 611-621, 1996.

ASTM, Standard Guide for Risk-Based Corrective Action Applied at Petroleum Release Sites (E1739-95), American Society for Testing and Materials, West Conshohocken, PA, 1995.

Atwood, D., and S. Gorelick, Hydraulic gradient control for groundwater contaminant removal, Journal of Hydrology, Amsterdam, The Netherlands, 76, 85-106, 1985.

Barker, J. F., G. C. Patrick, and D. Major, Natural Attenuation of Aromatic Hydrocarbons in a Shallow Sand Aquifer, *Ground Water Monitoring Review*, 7, 64-71, 1987.

Brainard, Jennifer, and David E. Burmaster, Bivariate Distributions for Height and Weight of Men and Women in the United States, *Risk Analysis*, 12(2), 1992.

Chang, L.-C., C. A. Shoemaker, and P. L.-F. Liu, Optimal time-varying pumping rates for groundwater remediation: application of a constrained optimal control algorithm, *Water Resources Research*, 28(12), 3157-3173, 1992.

Chang, Myung-Keun, Thomas C. Voice, and Craig S. Criddle, Kinetics of Competitive Inhibition and Cometabolism in the Biodegradation of Benzene, Toluene, and *p*-Xylene by Two Pseudomonas Isolates, *Biotechnology and Bioengineering*, 41(2), 1057-1065, 1993.

Culver, Teresa B., and Christine A. Shoemaker, Dynamic Optimal Control for Groundwater Remediation With Flexible Management Periods, *Water Resources Research*, 28(3), 629-641, 1992.

Davis, Andy, Susan Kamp, George Fennemore, Richard Schmidt, Michael Keating, Karl Hoenke, and Jeff Wyatt, A Risk-Based Approach to Soil Remediation Modeling *Environmental Science and Technology*, 31(11), 1997.

De Jong, K. A., An analysis of the behavior of a class of genetic adaptive systems, Ph.D. dissertation, University of Michigan, Ann Arbor, MI, 1975.

Deutsch, Clayton V., and Andre G. Journel, *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, New York, NY, 1992.

Domenico, Patrick A., and Franklin W. Schwartz, *Physical and Chemical Hydrogeology*, John Wiley and Sons, New York, NY, 1990.

Dosa, B., Testing, evaluation and sensitivity analysis of an in-situ bioremediation model, MS thesis, Cornell University, Ithaca, NY, 1994.

Dougherty, David E., and Robert A. Marryott, Optimal Groundwater Management; 1. Simulated Annealing, *Water Resources Research*, 27(10), 2493-2508, 1991.

Droppo, J. G., D. L. Strenge, J. W. Buck, B. L. Hoopes, R. D. Brockhaus, M. B. Walter, and G. Whelan, *Multimedia Environmental Pollutant Assessment System (MEPAS) Application Guidance Volume 2 – Guidelines for Evaluating MEPAS Parameters*, PNL-7216 (Pacific Northwest Laboratory, Richland, Washington, 1989.

Freeze, R. Allan, and John A. Cherry, *Groundwater*, Prentice Hall, Englewood Cliffs, NJ, 1979.

Gaudy, A. F., Jr., and Gaudy, E. T., *Microbiology for Environmental Scientists and Engineers*, McGraw-Hill, New York, NY, 1980.

Gelhar, Lynn W., *Stochastic Subsurface Hydrology*, Prentice Hall, Englewood Cliffs, NJ, 1993.

Goldberg, David E., *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, New York, NY, 1989.

Graham, Wendy D., and Dennis B. McLaughlin, A Stochastic Model of Solute Transport in Groundwater: Application to the Borden, Ontario, Tracer Test, *Water Resources Research*, 27(6), 1345-1359, 1991.

Gutjahr, Allan, Sean Hatch, and Bryan Bullard, Random Field Generation Using the Fast Fourier Transform, September 4, 1995.

Hoeksema, R. J., and P. K. Kitanidis, Analysis of the spatial structure of properties of selected aquifers, *Water Resources Research*, 21(4), 563-572, 1985.

Holland, J. H., *Adaptation in Natural and Artificial Systems*, University of Michigan, Ann Arbor, Michigan, 1975.

Hutchins, Stephen R., Biodegradation of Monoaromatic Hydrocarbons by Aquifer Microorganisms Using Oxygen, Nitrate, or Nitrous Oxide as the Terminal Electron Acceptor, *Applied and Environmental Microbiology*, 57(8), 2403-2407, 1991.

International Commission on Radiological Protection (ICRP), *Report of the Task Group on Reference Man*, ICRP Publication No. 23, Pergamon Press, New York, NY, 1975.

Israeli, Miron, and Christopher B. Nelson, Distribution and Expected Time of Residence for U.S. Households, *Risk Analysis*, 12(1), 1992.

James, I. R., and M. W. Knuiman, An Application of Bayes Methodology to the Analysis of Diary Records from a Water Use Study, *Journal of the American Statistical Association*, 82(399), 1987.

Journel, A. G., and Ch. J. Huijbregts, *Mining Geostatistics*, Academic Press, New York, NY, 1978.

Kangas, M., Probabilistic Risk Assessment, *ASTM Standardization News*, 24(6), 28-33, 1996.

Karatzas, George P., and George F. Pinder, Groundwater Management Using Numerical Simulation and the Outer Approximation Method for Global Optimization, *Water Resources Research*, 29(10), 3371-3378, 1993.

Karatzas, George P., and George F. Pinder, The solution of groundwater quality management problems with a nonconvex feasible region using a cutting plane optimization technique, *Water Resources Research*, 32(4), 1091-1100, 1996.

Labieniec, Paula A., David A. Dzombak, and Robert L. Siegrist, SoilRisk: Risk Assessment Model for Organic Contaminants in Soil, *Journal of Environmental Engineering*, May 1996a.

Labieniec, Paula A., David A. Dzombak, and Robert L. Siegrist, Risk Variability Due to Uniform Soil Remediation Goals, *Journal of Environmental Engineering*, July 1996b.

Lyman, Warren J., Patrick J. Reidy, and Benjamin Levy, *Mobility and Degradation of Organic Contaminants in Subsurface Environments*, C. K. Smoley, Chelsea, MI, 1992.

Mackay, D. M., D. L. Freyberg, P. V. Roberts, and J. A. Cherry, A Natural Gradient Experiment on Solute Transport in a Sand Aquifer; 1. Approach and Overview of Plume Movement, *Water Resources Research*, 22(13), 2017-2029, 1986.

Marryott, Robert A., and David E. Dougherty, and Robert L. Stollar, Optimal Groundwater Management; 2. Application of Simulated Annealing to a Field-Scale Contamination Site, *Water Resources Research*, 29(4), 847-860, 1993.

Marsily, Ghislain de, *Quantitative Hydrogeology*, Academic Press, New York, NY, 1986.

Maxwell, Reed M., Susan D. Pelmulder, Andrew F. B. Tompson, and William E. Kastenberg, On the development of a new methodology for groundwater-driven health risk management, *Water Resources Research*, 34(4), 833-847, 1998.

McKinney, D. C., and M.-D. Lin, Approximate mixed-integer nonlinear programming methods for optimal aquifer remediation design, *Water Resources Research*, 31(3) 731-740, 1995.

McKinney, D. C., and M.-D. Lin, Genetic algorithm solution of groundwater management models, *Water Resources Research*, 30(6), 1897-1906, 1994.

McKinney, D. C., and M.-D. Lin, Pump-and-Treat Ground-water Remediation System Optimization, *Journal of Water Resources Planning and Management*, March/April 1996.

McKone, Thomas E., Human Exposure to Volatile Organic Compounds in Household Tap Water: The Indoor Inhalation Pathway, *Environmental Science and Technology*, 21(12), 1987.

McKone, Thomas E., and Kenneth T. Bogen, Predicting the uncertainties in Risk Assessment: A California groundwater case study, *Environmental Science and Technology*, 25(10), 1991.

Miller, Brad, Noise, Sampling, and Efficient Genetic Algorithms, Ph.D. dissertation, University of Illinois at Urbana-Champaign, Urbana, IL, 1997.

Minsker, B. S., and C. A. Shoemaker, Dynamic Optimal Control of In-Situ Bioremediation of Ground Water, *Journal of Water Resources Planning and Management*, May/June 1998.

Morgan, David R., J. Wayland Eheart, and Albert J. Valocchi, Aquifer Remediation Design Under Uncertainty Using a New Chance Constrained Programming Technique, *Water Resources Research*, 29(3), 551-561, 1993.

Pelmulder, Susan D., and William W-G. Yeh, Regional scale framework for modeling water resources and health risk problems, *Water Resources Research*, 32(6), 1851-1861, 1996.

Rast, R., ed., *Environmental Restoration: Assemblies Cost Book*, R. S. Means Company, Kingston, MA, 1997a.

Rast, R., ed., *Environmental Restoration: Unit Cost Book*, R. S. Means Company, Kingston, MA, 1997b.

Ritzel, Brian J., J. Wayland Eheart, and S. Ranjithan, Using genetic algorithms to solve a multiple objective groundwater pollution containment problem, *Water Resources Research*, 30(5), 1589-1603, 1994.

Roseberry, Ann M., and David E. Burmaster, Lognormal Distributions for Water Intake by Children and Adults, *Risk Analysis*, 12(1), 1992.

Schaum, John, Kim Hoang, Russell Kinerson, Jackie Moya, and Rhoda G. M. Wang, Estimating Dermal and Inhalation Exposure to Volatile Chemicals in Domestic Water, *Water Contamination and Health: Integration of Exposure Assessment, Toxicology, and Risk Assessment*, Edited by Rhoda G. M. Wang, Marcel Dekker, New York, NY, 1994.

Smith, Roy L., Use of Monte Carlo Simulation for Human Exposure Assessment at a Superfund Site, *Risk Analysis*, 14(4), 1994.

Smith, A. E., P. B. Ryan, and J. S. Evans, The Effect of Neglecting Correlations When Propagating Uncertainty and Estimating the Population Distribution of Risk, *Risk Analysis*, 12(4), 1992.

Taylor, S. W., Modeling enhanced in-situ biodegradation in groundwater: Model response to biological parameter uncertainty, *Proceedings: 1993 Groundwater Modeling Conference*, International Ground Water Modeling Center, Golden, CO, 1993.

Tiedeman, Claire, and Steven M. Gorelick, Analysis of Uncertainty in Optimal Groundwater Contaminant Capture Design, *Water Resources Research*, 29(7), 2139-2153, 1993.

U.S. Environmental Protection Agency, *Integrated Risk Information System (IRIS) Substance File # 0276 - Benzene; CASRN 71-43-2 (04/01/97)*, On-line Database, 1997.

U.S. Environmental Protection Agency, *MMSOILS: Multimedia Contaminant Fate, Transport, and Exposure Model*, Office of Research and Development, Washington, DC, 1992.

U.S. Environmental Protection Agency, *Risk Assessment Guidance for Superfund Volume 1: Human Health Evaluation Manual (Part A), Interim Final*, Office of Emergency and Remedial Response (EPA/540/1-89/002), Washington, DC, 1989.

U.S. Environmental Protection Agency, *Risk Assessment Guidance for Superfund Volume 1: Human Health Evaluation Manual Supplemental Guidance, Standard Default Exposure Factors, Interim Final*, Office of Emergency and Remedial Response (OSWER Directive 9285.6-03), Washington, DC, 1991.

U.S. Environmental Protection Agency, *Superfund Exposure Assessment Manual*, Office of Emergency and Remedial Response (EPA/540/1-88/001; OSWER Directive 9285.5-1), Washington, DC, 1988.

U.S. Environmental Protection Agency, *Use of Monte Carlo Simulation in Risk Assessments*, Region 3, Hazardous Waste Management Division, Office of Superfund Programs (EPA903-F-94-001), Philadelphia, PA, 1994.

Wagner, B. J., and S. M. Gorelick, Optimal Groundwater Quality Management Under Parameter Uncertainty, *Water Resources Research*, 23(7), 1162-1174, 1987.

Wagner, B. J., and S. M. Gorelick, Reliable Aquifer Remediation in the Presence of Spatially Variable Hydraulic Conductivity: From Data to Design, *Water Resources Research*, 25(10), 2211-2225, 1989.