Air Force Institute of Technology

AFIT Scholar

3-2000

Use of Genetic Algorithms to Characterize Groundwater Contamination Source Areas

Chaz M. Williamson

Follow this and additional works at: https://scholar.afit.edu/etd

Part of the Environmental Engineering Commons

USE OF GENETIC ALGORITHMS TO
CHARACTERIZE GROUNDWATER
CONTAMINATION SOURCE AREAS

THESIS

Chaz M. Williamson, Captain, USAF

AFIT/GEE/ENV/00M-17

20000515 015

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 2000	Master's Thesis

4. TITLE AND SUBTITLE
USE OF GENETIC ALGORITHMS TO CHARACTERIZE GROUNDWATER CONTAMINATION SOURCE AREAS

5. FUNDING NUMBERS

6. AUTHOR(S)
Chaz M. Williamson, Captain, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
2750 P Street
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GEE/ENV/00M-17

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFRL/MLQ
139 Barnes Drive, Suite 2
Tyndall AFB FL 32403-5323

Ms. Alison Lightner
(850) 283-6303

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Dr. Mark N. Goltz, ENV

(937) 255-3636 x4638
mark.goltz@afit.af.mil

12a. DISTRIBUTION AVAILABILITY STATEMENT

Approved for public release, distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)
In this work, genetic algorithms (GAs) were used to help interpret tracer breakthrough curves from partitioning interwell tracer tests (PITTs) conducted at Hill AFB, Utah by researchers from the University of Florida. Two transport models were developed to simulate tracer transport in the test cells. One model assumed the cell consisted of multiple layers, and that transport in each layer could be described by the one-dimensional advective/dispersive equation. The second model also assumed multiple layers, and modeled transport in the individual layes as advective transport through 100 tubes. Transport times were represented by a stochastic (lognormal) distribution. The model solutions were coded into Microsoft Excel. Model parameters were optimized using Evolutionary Solver, a GA developed by Frontline Systems. The optimized parameters were used to estimate pre-and post-flushing NAPL saturations, as well as cleanup efficiency. Results were compared to estimates obtained through moment analysis of the PITT data. Results demonstrated that GAs are a tool that may be useful in interpreting PITT data for the characterization of NAPL source areas. In particular, using the GAs to interpret PITT data provided more information than could be obtained from moment analysis.

14. SUBJECT TERMS
genetic algorithm (GA), partitioning interwell tracer test (PITT), non-aqueous phase liquid (NAPL), inverse modeling, groundwater

15. NUMBER OF PAGES
86

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

The views expressed in this thesis are those of the author and do not reflect

the official policy or position of the United States Air Force,

Department of Defense or U.S. Government.

USE OF GENETIC ALGORITHMS TO CHARACTERIZE

GROUNDWATER CONTAMINATION SOURCE AREAS

THESIS

Presented to the Faculty of the Graduate School of Engineering and Management

of the Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Engineering and Environmental Management

Chaz M. Williamson, B.S.
Captain, USAF
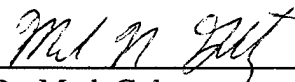
March 2000

USE OF GENETIC ALGORITHMS TO CHARACTERIZE
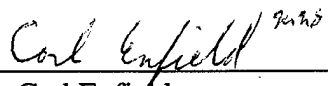
GROUNDWATER CONTAMINATION SOURCE AREAS

Chaz M. Williamson, B.S.
Captain, USAF
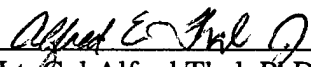
Approved:
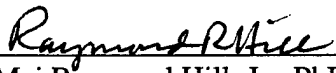
_____
Dr. Mark Goltz
Co-Chairman

_____
Dr. Carl Enfield
Co-Chairman

_____
Lt/Col Alfred Thal, PhD
Member

_____
Maj Raymond Hill, Jr., PhD
Member

Acknowledgements

I would like to thank my co-advisors, Dr. Mark Goltz and Dr. Carl Enfield, for willingly providing their knowledge and expertise to this effort. They both demonstrated remarkable patience in answering my many questions, some of which I repeated several times. Maj Ray Hill and Lt Col Al Thal, who served as committee members, also provided valuable assistance in completing this work. Maj Hill's knowledge of Excel and mathematical modeling provided a critical foundation to building the groundwater flow models used in this thesis. Lt Col Thal's first-hand knowledge of the tests conducted at Hill AFB saved me hours of frustration in interpreting the data.

I would especially like to thank my wife, Eileen, for her patience over the past several months as I worked to finish this thesis. I would also like to thank her for freely giving her time and skills to provide many of the graphics included in the document.

Table of Contents

List of Figures

List of Tables

Abstract

This research investigates the application of genetic algorithms (GAs) to help interpret data from partitioning interwell tracer tests (PITTs) to characterize groundwater contamination source areas. The data used in this research were obtained from PITTs conducted in hydraulically isolated test cells at Hill AFB, Utah by researchers from the University of Florida. The tests were carried out to evaluate the effectiveness of cosolvent and surfactant flushing for remediating non-aqueous phase liquid (NAPL) sources of groundwater contamination. PITTs use tracers that flow from an injection to an extraction well in the test cells. The quantity and distribution of NAPL in the cell can be inferred by the tracer concentration versus time responses (known as the breakthrough curves) at the extraction well.

In this work, GAs were used to help interpret tracer breakthrough curves from PITTs. Two transport models were developed to simulate tracer transport in the test cells. One model assumed the cell consisted of multiple layers, and that transport in each layer could be described by the one-dimensional advective/dispersive equation. The second model also assumed multiple layers, and modeled transport in the individual layers as advective transport through 100 tubes. Transport times through the tubes were represented by a stochastic (lognormal) distribution. The model solutions were coded into Microsoft Excel. Model parameters were optimized using Evolutionary Solver, a GA developed by Frontline Systems. The optimized parameters were used to estimate pre- and post-flushing NAPL saturation, as well as cleanup efficiency. Results were compared to estimates obtained through moment analysis of the PITT data. Results demonstrated that GAs are a tool that may be useful in interpreting PITT data for the

characterization of NAPL source areas. In particular, using the GAs to interpret the PITT

data provided more information on NAPL distribution than could be obtained from

moment analysis.

USE OF GENETIC ALGORITHMS TO CHARACTERIZE

GROUNDWATER CONTAMINATION SOURCE AREAS

Chapter 1: Introduction

1.1 Background

Non-aqueous phase liquids (NAPLs) encompass a broad range of contaminants,
including petroleum products and solvents, with the shared characteristic of a low
solubility in water. NAPLs are further divided into dense NAPLs (DNAPLs) that have a
higher density than water, and light NAPLs (LNAPLs) that have a lower density than
water. NAPL contaminants migrate downward through the vadose zone because of
gravity and capillary forces. Once the contaminants reach the water table, DNAPLs tend
to migrate downward through the saturated aquifer and may pool on low permeability
soil layers while LNAPLs tend to pool on top of the water table. In both cases, NAPL
contamination of groundwater may involve a separate phase NAPL source that slowly
dissolves into the groundwater to form an aqueous-phase plume. Such sources can
remain for long time periods, ranging from several decades to centuries, and no
remediation method has been demonstrated to effectively restore sites contaminated by
NAPLs (Trowbridge *et al.*, 1999). Because of widespread applications of solvents at
military facilities and past disposal practices that resulted in contamination by these
solvents, the problem of NAPL remediation is a significant concern for the Department of
Defense (DOD) and the Air Force (Armstrong Laboratory, 1997).

Tests involving several innovative remediation methods have shown encouraging results for dealing with the NAPL problem. Recent tests involving cosolvent and surfactant flushing (Lowe *et al.*, 1999; Rao *et al.*, 1997; Jawitz *et al.*, 1998, Falta *et al.*, 1999) demonstrate that this emerging technology may offer a potential remedy for dealing with NAPL source areas. Cosolvent (surfactant) flushing involves the injection of a cosolvent (surfactant) / water mixture near the source area of a contaminated site to solubilize or mobilize the contaminant. This allows subsequent extraction of the contaminant from the subsurface for aboveground treatment. Field applications of cosolvent (surfactant) floods have demonstrated the ability of the technology to remove significant amounts of NAPL contaminant. In 1996, field trials involving the cosolvent flushing of a NAPL-contaminated site at Hill AFB, Utah, recovered between 70 – 90% of petroleum hydrocarbons and spent solvents (Lowe *et al.*, 1999).

Proper characterization of the NAPL contaminant source area is necessary to develop an effective remediation method and includes determining the location, composition, and quantity of NAPL at the site. Current methods for characterizing the NAPL contamination at a site include core sampling, cone penetrometer testing, geophysical logging, and partitioning inter-well tracer tests (PITT) (Jin *et al.*, 1995). Sections 2.1.4 and 2.3.1 discuss the benefits and limitations of each of these methods. The PITT method, as outlined by Jin *et al.* (1995), involves the simultaneous injection of a conservative (or non-partitioning) tracer and tracers that partition into the NAPL contaminant. The conservative and partitioning tracers are subsequently recovered at an extraction well. Partitioning of the non-conservative tracers into the separate phase NAPL retards the tracers' respective velocities. Several partitioning tracers are used so

2

that data can be obtained in a reasonable amount of time with adequate separation of the breakthrough curves for the conservative and partitioning tracers. In general, the appropriate tracer will have a retardation factor between 1.2 and 4, where the retardation factor represents the ratio of the groundwater velocity to the velocity of the partitioning tracer (Jin *et al.*, 1997). Since the retardation factor of the partitioning tracer is a function of the average NAPL saturation (i.e., a higher retardation factor results from a higher NAPL saturation), the breakthrough curves can be used to estimate the mass of separate phase NAPL (Young *et al.*, 1999). Soil composition, particularly the fraction of organic carbon, can affect the retardation of the partitioning tracer. Therefore, soil composition must be considered in the PITT analysis, and the PITT method may not be appropriate for soils with a high fraction of organic carbon.

Since the PITT method samples a much larger volume of the aquifer, it may provide a better overall characterization of the NAPL than the other methods (Jin *et al.*, 1995). Field evaluations performed at Operable Unit 1 (OU1), Hill AFB, Utah in 1995 (Falta *et al.*, 1999) utilized the PITT method, performing pre-flood and post-flood tracer tests, to estimate the efficiency of cosolvent flooding on an LNAPL contaminant. Results from these tests indicated removal of approximately 78% of the bulk NAPL from the test cell.

Natural aquifers typically have a high degree of variability in soil characteristics, such as hydraulic conductivity, and even aquifers characterized as homogenous may be considerably heterogeneous. The degree of soil heterogeneity at a site may be the most important factor affecting NAPL distribution and the amount of cosolvent (surfactant) / water mixture that is able to contact the contaminant (Lowe *et al.*, 1999). As the NAPL

phase is typically non-wetting, the mass of NAPL able to penetrate a pore space decreases at a disproportionate rate with a decrease in pore size. Consequently, the NAPL tends to accumulate in the larger pore spaces and a lower surface area to volume ratio is available for interaction with the cosolvent flood. In addition, the cosolvent (surfactant) / water mixture will preferentially flow through regions of high permeability. Since permeability is directly related to pore size, the preferential flow of the cosolvent (surfactant) / water mixture through the high permeability regions will tend to bring more of the mixture into contact with the bulk of the NAPL mass. However, it is also less likely that the mixture will contact the NAPL mass occupying smaller pore spaces. These behaviors of the NAPL and cosolvent (surfactant) / water mixtures imply that hydraulic conductivity distributions within the porous media are important factors to consider when designing the remediation method and calculating the quantity of cosolvent mixture required. The ability to develop a useful model of these distributions is a critical step in designing an effective cosolvent flushing system.

1.2 Scope of Research

The goal of this research is to develop a modeling approach to better predict residual NAPL saturations and distribution at a contaminated site (using data from pre-flood PITT tests) in an attempt to improve remediation design and better predict cleanup efficiency. This research will address the following questions:

> 1. In developing the model, and in deciding which modeling approach best accomplishes the purposes of this thesis, what objective function(s) should be optimized? Should the model minimize sum of squares difference between modeled and actual breakthrough data, minimize the number of fitting parameters, or some combination of the two objectives? What optimization method should be used to determine the best-fit model parameters?

2. Should only data from the non-partitioning tracer be used to determine the hydraulic conductivity and other groundwater flow parameters, or should the partitioning tracer data also be used? How should the partitioning tracer data be used to determine NAPL distribution? Do the models demonstrate a relationship between the hydraulic conductivity and NAPL distributions at the site?

3. How can the models be used to predict cleanup efficiency? How do model predictions compare to field data?

Several models reported in the literature have been developed to model groundwater transport. This research will use two modeling approaches to investigate the applicability of Genetic Algorithms (GAs) as an optimization method to estimate parameter values for interpretation of PITT results. One approach uses a model based on the analytic solution to the one-dimensional advective / dispersive equation (van Genuchten and Alves, 1982). The other approach uses a model based on a representation of the breakthrough curve using a stochastic function (Enfield, 2000). The models will be applied to data obtained from PITTs performed at OU1.

Chapter 2. Literature Review

2.1 Non-Aqueous Phase Liquids (NAPLs)

NAPLs are characterized by their low solubility in water. This property gives them the potential to exist as a separate phase both above and below the water table. NAPLs may be placed in three broad categories based on the specific gravity of the contaminant. Light NAPLs (LNAPLs) are less dense than water and tend to spread out on top of the capillary fringe at the transition from the vadose to saturated zones. Dense NAPLs (DNAPLs) are more dense than water and able to migrate downward through the saturated zone. Neutrally Buoyant NAPLs (NNAPLs), usually a mixture of LNAPLs and DNAPLs, have a specific gravity near that of water and vertical migration through the groundwater occurs slowly.

2.1.1 NAPL Properties

Commonly, a variety of components comprise the NAPL at a given site, and the characteristics of the NAPL contaminant are a composite of the individual properties of the components. These properties include density, viscosity, wettability, interfacial tension, and capillary pressure. *Density*, defined as mass per unit volume, and the specific gravity (density relative to the density of water) of the NAPL significantly impact its behavior within the saturated zone. *Viscosity* is a measure of a fluid's resistance to flow, and the viscosity of the NAPL defines its ability to flow as a separate phase. In general, a NAPL with a viscosity less than water will be highly mobile in the subsurface while a NAPL that is significantly more viscous than water will be practically immobile as a separate phase. *Wettability* refers to the affinity of a liquid for a solid surface in the presence of another immiscible liquid(s), and is determined by the angle of

the interface between the solid and the liquids. The wetting phase is the fluid through which the angle is less than 90 degrees, while the non-wetting phase has an angle greater than 90 degrees. Because they are non-polar molecules, NAPLs are non-wetting in the presence of water (although high dissolved organic content in the water can affect wettability) and wetting in the presence of air. *Interfacial tension* describes the tensile forces acting on the interface between fluids, and a high interfacial tension indicates a low affinity between the two liquids.

In the case of an immiscible liquid contaminant in groundwater, high interfacial tension results in the immiscible liquid molecules grouping in a way that minimizes the interfacial area. An interfacial tension of zero indicates miscibility, such that the liquid has entered the aqueous, or dissolved, phase. The interfacial tension at the interface of two immiscible liquids results in a pressure differential across that interface. This pressure differential, directly proportional to the interfacial tension, is referred to as *capillary pressure*. The pressure required for the non-wetting phase to enter a region saturated by the wetting phase (i.e., water), referred to as the *displacement pressure*, must be greater than the capillary pressure. The drying curve in Figure 1 illustrates that as the non-wetting phase occupies more of the pore space, the capillary pressure increases until the non-wetting phase reaches maximum saturation. As the wetting phase re-enters the pore space (wetting curve, Figure 1), the non-wetting phase saturation decreases until it reaches the point at which flow of the non-wetting phase is no longer possible. This *residual saturation* represents a non-flowing separate phase NAPL source that causes long-term contamination as it dissolves into the flowing groundwater.

7
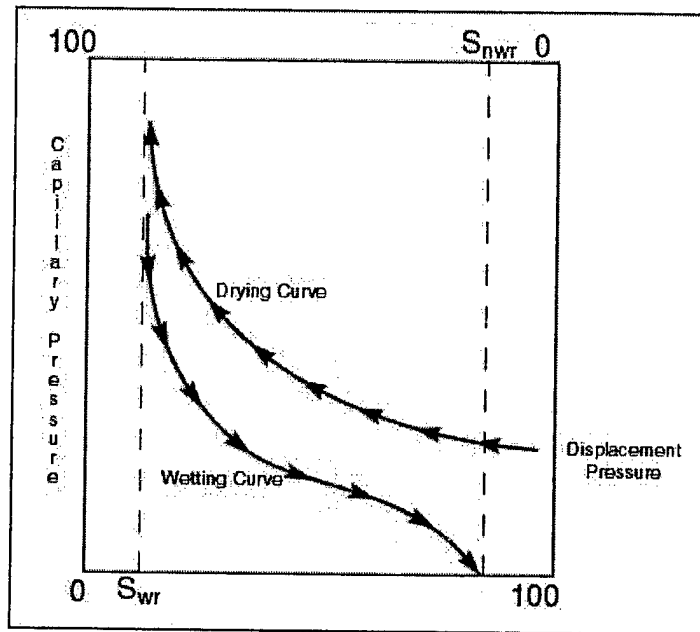
Figure 1: Wetting and Drying Curves (Domenico and Schwartz, 1998)

2.1.2 NAPL Phases

NAPLs are generally present at a contaminated site in four phases: as vapor phase, dissolved (or aqueous) phase, separate phase, and sorbed phase. NAPLs in the vapor phase may be removed by a variety of techniques including soil-vapor extraction (SVE) from the vadose zone. Air sparging, the injection and recovery of air into the saturated zone, can be used to recover dissolved components of volatile NAPLs. Active remediation methods such as pump-and-treat or passive methods such as funnel and gate can remove dissolved phase contaminant. Separate phase contaminant may be present as a residual or pooled NAPL. Depending on the viscosity of the pooled, or free-phase NAPL, the contaminant may be mobile in the subsurface and may be capable of being removed by direct pumping. LNAPLs can typically be found as free-phase contaminant

in the vadose zone and near the capillary fringe, where fluctuations in the depth of the water table result in NAPL at residual saturation in a smear zone both above and below the water table. DNAPLs, on the other hand, migrate downward through the saturated zone until they encounter a low permeability barrier. The discovery of free-phase DNAPL pools containing significant mass is rare, and the presence of DNAPLs is typically inferred from site history or groundwater monitoring data (Sellers, 1999).

2.1.3 NAPL Migration in the Subsurface

The properties of DNAPLs and LNAPLs cause differences in their distribution and residual saturation at a contaminated site. The LNAPL smear zone generally contains uniform distribution of LNAPL at residual saturation. DNAPLs tend to migrate downward through the saturated zone by way of preferential flow paths. When low permeability layers are encountered, the DNAPL spreads horizontally until it encounters higher permeability soil and vertical migration resumes. Vertical migration may also resume when the depth of pooled NAPL becomes such that it can no longer be supported by hydrostatic pressure (entry pressure) of the low permeability region. The resulting "lenses" of pooled NAPL and "fingers" of residual NAPL can be difficult to locate, and DNAPL saturation levels can have a high degree of variability in the saturated zone. Figure 2 shows a conceptual diagram of LNAPL and DNAPL contamination of the saturated zone.
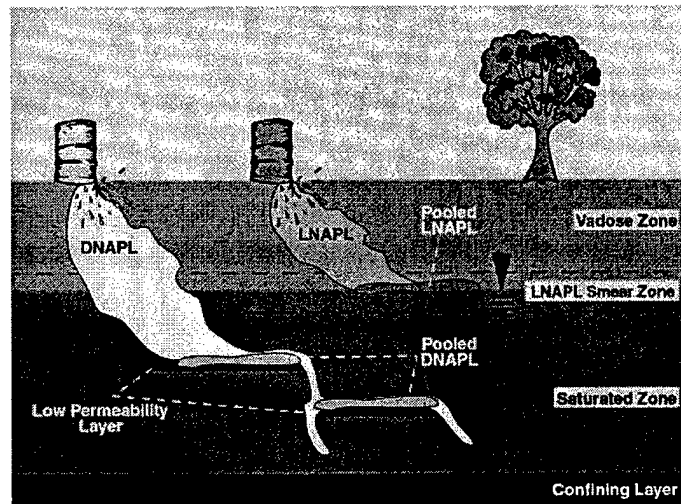
Figure 2: NAPL Distribution at a Contaminated Site

Experiments by Broholm *et al.* (1999) investigated the ability to estimate DNAPL

source distribution through down-gradient monitoring of the aqueous phase plume. In

their experiments, results obtained from detailed groundwater monitoring were compared

to the DNAPL source distribution found by collecting core samples from the source area.

The results demonstrated a general correlation between the spatial delineation of the

down-gradient plume to the vertical and lateral distribution of the source. However,

detailed groundwater monitoring is typically insufficient to determine small-scale

distribution of the DNAPL source. Detailed core sampling was able to account for only

67% - 87% of the known mass injected into the test cell. These tests demonstrated the

difficulty in obtaining an accurate estimate of NAPL mass and distribution, even in a

controlled experiment.

The presence of two or more fluids within a pore space results in reduction in permeability to any one of those fluids. As the saturation of one fluid increases, the permeability of the pore space to another fluid decreases (Lowe *et al.*, 1999). Because of this, the distribution of NAPL in the saturated zone impacts the hydraulic characteristics of an aquifer, and the presence of residual NAPL in the pore spaces can significantly reduce hydraulic conductivity. This effect can impede the ability of mobility enhancing agents, such as cosolvents and surfactants, to interact with the separate phase contaminant. Therefore, proper characterization of contaminant distribution is critical prior to developing a remediation strategy. However, technological and financial constraints limit the ability to develop a clear picture of NAPL distribution in the source area. Obviously, some characterization of the source area must be obtained, but the level of detail required remains a subject of discussion.

2.1.4 Characterizing a NAPL Source Area

There are four commonly used methods to characterize a NAPL source area: core sampling, cone penetrometry, geophysical logging, and partitioning inter-well tracer tests (PITTs). The first three methods involve direct examination of the source area. Core samples allow visual examination of the NAPL distribution in the saturated zone, can be used to estimate the boundaries of the source area, and can be extrapolated using Kriging analysis to estimate total NAPL mass. Core samples provide a valuable tool for source area characterization, but are generally inadequate for estimating total NAPL mass because of the wide variability in NAPL (especially DNAPL) distribution discussed in Section 2.1.2 (Broholm *et al.*, 1999). Cone penetrometry provides similar data to that obtained from core samples, although data are obtained without the need to extract a core.

11

Typically, cone penetrometry utilizes laser-induced fluorescence (LIF) to identify NAPL in the subsurface. LIF is based on the principle that aromatic hydrocarbons fluoresce when contacted by a laser. Intrusive characterization methods such as core sampling and cone penetrometry have the drawback that disturbance of the subsurface could cause vertical mobilization of pooled DNAPL. Geophysical logging can be used to deduct the distribution of NAPL from subterranean characteristics, but is inadequate for locating NAPL mass. Unlike the first three methods that examine the source area directly, PITTs characterizes the source zone indirectly using partitioning and conservative tracers. Downgradient monitoring of aqueous phase contaminant concentrations can also be used to characterize the vertical distribution of the NAPL (Broholm *et al.*, 1999).

2.1.5 NAPL Remediation

There are four general objectives for groundwater remediation (Sellers, 1999): prevent exposure to contaminated groundwater, contain the contaminant plume, control or reduce the source mass, and finally restore the aquifer to the greatest extent practicable. The desirability of the remediation objective increases from preventing exposure to restoration, but attainability of objectives is limited by economic and technological constraints. Preventing exposure may involve providing an alternative water source to property owners affected by the contaminated groundwater. In the case of NAPL-contaminated groundwater, containment involves controlling the flow of the aqueous phase plume and in-situ or above ground removal of the aqueous phase contaminant. The hydrophobic properties of NAPLs, along with the variability in separate phase distribution, typically limit our ability to reduce the source mass and restore the contaminated aquifer.

Since traditional remediation methods such as pump and treat have proven to be ineffective at reducing the NAPL source mass, the strategy typically taken at sites with a NAPL source area has been to contain the aqueous phase plume, often using "pump and treat" extraction wells. Other source area and plume containment methods such as sheet piling, slurry walls, and funnel and gate systems have been used with some success, but all of these methods have high long-term costs and do nothing to reduce the source mass.

Because of the inadequacy of pump and treat remediation of NAPL-contaminated groundwater, the EPA has suggested a "Triple Train Response" (Mercer, 1991) as a possible remediation approach. In this three-step process, the first step would be to install extraction wells screened at an appropriate depth to remove mobile separate phase contaminant. The second step would use techniques to remove contaminant at residual saturation. In the final step, pump and treat would be used to remediate the aqueous phase plume. A field test of the Triple Train approach at a Superfund site in Laramie, Wyoming, removed 99.8% of the contaminant mass (Mercer, 1991). Emerging technologies such as *cosolvent* and *surfactant* that act directly on the separate phase contaminant would be applied at the second step of the triple train approach to reduce source mass. Other approaches for reducing source mass include steam injection and air sparging. Steam injection mobilizes the separate phase NAPL by lowering its viscosity, increasing its volatility, and inducing a hydraulic gradient that can mobilize the contaminant (Sellers, 1999). Air sparging, which involves the injection of air into the saturated zone, can be an effective source reduction method for volatile contaminants.

2.2 Cosolvent / Surfactant Flushing

Laboratory and field tests involving cosolvents and surfactants have demonstrated

the potential of the technologies as cost effective alternatives for remediation of

contaminated sites (Anason, 1999). Cosolvent or surfactant flushing involves the

injection of a cosolvent (surfactant) / water mixture, commonly called the cosolvent

(surfactant) flood, near the NAPL source area. Through the processes of solubilization

and mobilization, detailed in Section 2.2.1, a significant portion of the NAPL source

mass is removed for subsequent above ground treatment in a relatively short period of

time.

2.2.1 Solubilization and Mobilization

Surfactants, or surface active agents, affect the interface between the separate

phase NAPL and the water. Surfactants are typically organic compounds with long

hydrophobic non-polar carbon chains with a strongly polar hydrophilic end. Surfactant

floods are often applied as a mixture of surfactants and co-surfactants (typically

intermediate chain alcohols that act as surfactants in the presence of other surfactants) to

enhance effectiveness. The imbedding of the hydrophobic end into the NAPL reduces

the NAPL / water interfacial tension and may mobilize the contaminant. The formation

of micelles, conglomerations of surfactant that form a separate hydrophophic phase into

which the NAPL molecules can partition, increases the solubility of the NAPL in the

water. This process is referred to as micellar solubilization, and occurs at a threshold

surfactant concentration called the critical micelle concentration. Surfactants also form

micro-emulsions, suspensions of microscopic droplets of one immiscible liquid in

another immiscible liquid, in water. Surfactants include common detergents, and,

although only recently applied to groundwater remediation, have long been used by the oil industry for enhanced oil recovery (Lowe *et al.*, 1999).

Surfactant mixtures are classified according to the type of micro-emulsion that forms. In Winsor Type I systems, NAPL droplets form in a continuous water phase. In a Winsor Type II system, water droplets form in a continuous oil phase and are used when the goal is mobilization of the contaminant. Winsor Type III systems, or middle-phase systems, fall somewhere between the Type I and Type II systems, and result in lower interfacial tensions than can be achieved in either of the other two systems. Because of this, Type III systems result in solubilization of the contaminant. However, since Type III systems require the optimization of a large number of parameters, they are difficult to apply in the field (Jawitz *et al.*, 1998).

Cosolvents are organic compounds, typically an alcohol, that are miscible in both water and NAPL. Cosolvents may be used with surfactants to enhance surfactant performance or on their own to increase dissolution or induce mobilization of the NAPL. When used at low concentrations, cosolvents increase the aqueous solubility of many organic contaminants. At higher concentrations, cosolvents may partition into both the water and NAPL phases, reducing the interfacial tension and viscosity of the NAPL to the point that mobilization may occur. If sufficient quantity of cosolvent / water mixture is present, the NAPL may solubilize solely into the cosolvent / water mixture (Lowe *et al.*, 1999). Because organic compounds readily dissolve into organic compounds, NAPLs may enter the cosolvent flood in the aqueous phase and are able to be recovered for subsequent treatment.

15

As noted above, cosolvents and surfactants can enhance migration of separate phase NAPL in the saturated zone by two methods: solubilization or mobilization. Cosolvents or surfactants acting on the surface of the separate phase contaminant reduce the interfacial tension. In the case of surfactants, if a critical number of surfactant molecules are present, micelles form. NAPL is then able to partition into the micelle phase. Since displacement pressure is proportional to interfacial tension, the interfacial tension may be lowered to a point that mobilizes the NAPL. In this case, separate phase flow of the NAPL may occur. In some cases, mobilization may be the desired result, but in the case of DNAPLs, mobilization may result in loss of hydraulic control in which the separate phase NAPL does not flow with the flood.

2.2.2 Concerns

Because the solvents and surfactants used in a cosolvent flood may themselves be contaminants with regulatory maximum contaminant levels (MCLs), their use may be restricted and total recovery of the solvents after injection may be a critical concern. Another concern is the high cost resulting from the large quantity of solvent required to solubilize the separate phase NAPL. Field experiments have been conducted on recycling the solvent following aboveground treatment, but dissolved contaminants must be removed before re-injection. A variety of systems exist to purify and reuse the cosolvent / surfactant flood. Such systems are expensive to operate and have variable results (Anason, 1999). Additional problems can result from unstable flow conditions caused by density and viscosity differences between the cosolvent flood and groundwater (Armstrong Laboratory, 1999). Because of the potential for separate phase mobilization of DNAPLs, hydraulic control must be maintained. Another concern is uniform delivery

16

of the cosolvent to the source area to prevent partial removal or concentration of the contaminant. The NAPL source area and the hydraulic properties of the aquifer must be properly characterized to ensure interaction of the cosolvent or surfactant with the contaminant.

2.3 Partitioning Inter-well Tracer Tests (PITTs)

Partitioning tracer tests, first used by the petroleum industry in the early 1970s, take advantage of the chemical interaction between the separate phase organic compounds and tracers with an affinity towards those compounds. Four mechanisms result in retardation, or a reduced transport velocity, of the partitioning tracers: fluid partitioning, adsorption, ion exchange, and size exclusion. Fluid partitioning represents the most significant mechanism for tracer retardation when NAPL is present (Tang, 1995). When a partitioning and non-partitioning (conservative) tracer are injected simultaneously into NAPL-contaminated groundwater and subsequently recovered, the difference in their transport velocity (as evidenced by the separation of their breakthrough curves) can be used to estimate the mass of organic present in the aquifer.

Jin *et al.* (1995) first demonstrated PITTs in the field to characterize the NAPL saturation at sites with groundwater contamination and to assess remediation performance. Prior to conducting the test, the general location and dimensions of the source area must be determined, typically by using core sampling, cone penetrometry, or geologic mapping. The PITT involves the simultaneous injection of partitioning and conservative tracers at an injection well located up-gradient from the separate phase NAPL source area (Figure 3). The tracers are subsequently recovered at an extraction well located down-gradient from the separate phase NAPL. The NAPL saturation at the

17

site can be estimated from the separation of the breakthrough curves of the partitioning

and conservative tracers (Figures 4).
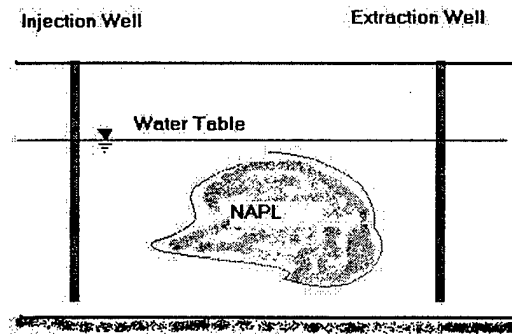
Figure 3: General PITT Test Configuration
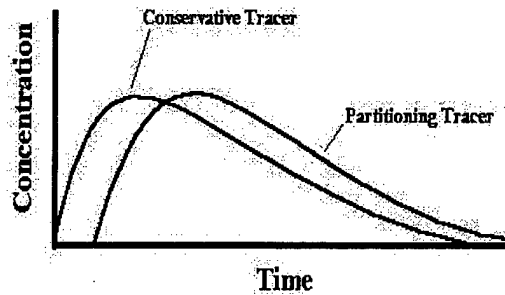
Figure 4: Idealized Breakthrough Curves

2.3.1 Limitations of PITTs

As with any method of NAPL site characterization effort, there are limitations on

the ability of PITTs to provide an accurate estimate of NAPL mass and cleanup

efficiency. Research has shown that partitioning tracers tend to underestimate the

saturation of NAPL in an aquifer (Nelson *et al.*, 1999). The presence of NAPL at

18

saturation in the pore spaces results in the reduction of the effective permeability of the

region occupied by the NAPL. The preferential flow of the tracer around these regions of

reduced hydraulic conductivity provides a likely explanation for the underestimation of

NAPL mass. Rate limited mass transfer and mass loss in the tracer also contribute to the

underestimation of NAPL mass (Nelson and Brusseau, 1996) Preferential tracer flow

may result in the tracers following a similar flow path as the cosolvent / surfactant flood.

Consequently, PITTs also tend to overestimate the cleanup efficiency of the flood.

Because of the limitations associated with any characterization method, two independent

methods (e.g., PITTs and core samples) are generally used to estimate NAPL mass and

cleanup efficiency (Lowe *et al.*, 1999).

2.3.2 Estimating NAPL Mass and Cleanup Efficiency from PITT Data

This research utilizes two methods for estimating NAPL saturation and cleanup

efficiency from PITT data: the method of moments and inverse modeling.

2.3.2.1 Method of Moments

Retardation (R) can be defined as the ratio of the transport velocity of the

partitioning tracer (v_p) over the conservative tracer (v_c) and can be calculated by (Sheely

and Baldwin, 1982),

$$R = 1 + \frac{K_{nw}S_n}{1 - S_n} = \frac{v_c}{v_p} = \frac{t_p}{t_c} \tag{1}$$

where K_{nw} is the partition coefficient of the tracer between the NAPL and water,

S_n is the NAPL saturation,

t_p is the travel time of the partitioning tracer,

and t_c is the travel time of the conservative tracer.

Rewriting (1) in terms of NAPL saturation yields,

19

$$S_n = \frac{t_p - t_n}{t_p + t_n(K_{nw} - 1)} \cdot \quad (2)$$

The mean tracer travel times and the standard deviation for the travel time

distribution can be estimated directly from the experimental data using the method of

moments. For a continuous random variable, the expected (E) or mean (μ) value of a

random variable X with a probability distribution function f(x) is defined as,

$$\mu_X = E(X) = \int_{-\infty}^{\infty} xf(x)dx \quad (3)$$

The variance (σ^2) of X is defined as,

$$\sigma^2 = E(X^2) - [E(X)]^2 \quad (4)$$

The first moment ($\mu'_{1,t}$) of the experimental data can be calculated using (3)

normalized to the data.

$$\mu'_{1,t} = \frac{\int_0^\infty tC(x,t)dt}{\int_0^\infty C(x,t)dt} \quad (5)$$

Adjusting (5) for tracer injection time yields the equation for mean travel time for the

tracers.

$$t_i = \frac{\int_0^\infty C_i t \, dt}{\int_0^\infty C_i dt} - \frac{t_0}{2} \quad (6)$$

where C_i is the tracer concentration at time t,

   t_i is the mean tracer travel time,

and t_0 is the tracer injection time.

The absolute second moment ($\mu'_{2,t}$), equivalent to $E(X^2)$ in (4), is calculated using,

$$\mu'_{2,t} = \frac{\int_0^\infty t^2 C(x,t)dt}{\int_0^\infty C(x,t)dt} \qquad (7)$$

The variance for the mean travel time can be calculated by substituting (5) and (7) into (4) to obtain,

$$\sigma^2 = \mu'_{2,t} - (\mu'_{1,t})^2 \qquad (8)$$

2.3.2.2 Inverse Modeling

Inverse modeling involves determining a function of *decision variables* that provides a reasonable representation of experimental data. There are a variety of approaches to optimize variable values and determine the best model fit to the experimental data by minimizing an *objective function* that represents the difference between the model and experimental data. Most optimization approaches utilize a variation of linear programming in which restraints are placed on decision variables to represent them as linear functions. Another approach, *combinatorial optimization*, represents the decision variables as discrete values and determines the optimum combination of those discrete values (Reeves, 1993). The two general classes of approaches to optimization problems described above can be classified as derivative based methods and search methods respectively (Lybanon and Messa, 1999).

One characteristic of combinatorial optimization is the existence of many *locally optimum* solutions. In other words, solution sets exist that may be more optimum than adjacent solution sets. Consequently, convergence may occur at a solution set that does not represent the optimum of all solution sets. Combinatorial optimization methods

21

based on linear programming use exact methods to guarantee convergence at a global optimum for the objective function (Reeves, 1993). However, because decision variables are constrained by linear or non-linear functions, convergence at the global optimum for the objective function does not necessarily guarantee that the results represent the best possible solution set for the model. A *heuristic* approach, on the other hand, utilizes a "seeking" method (Reeves, 1993) that does not guarantee convergence at a global optimum. It is possible, however, that iterative optimization using a heuristic approach can provide a good solution quicker than an approach that guarantees finding a global optimum for the objective function. Genetic algorithms (GAs), discussed in detail in Section 2.4, are an example of a heuristic approach to a combinatorial optimization problem. This research evaluates the applicability of GAs to groundwater transport modeling.

A variety of mathematical models have been developed for groundwater flow and remediation problems. Although these models may provide a reasonable simulation of groundwater flow and contaminant transport, site complexity often prevents their application due to the difficulty and expense associated with data collection and parameter quantification (Wang, 1997).

2.3.2.2.1 Analytic Modeling Approach

Analytic solutions to the differential equations that describe groundwater transport provide one modeling approach. The simplifying assumption of steady one-dimensional flow will be used in this research to develop the chemical transport model. The assumption of one-dimensional flow is generally inadequate for a three-dimensional flow problem and would be expected to provide reasonable results only for homogenous

conditions that would typically not exist in the field. Therefore, a close correlation of the model results to the field data is not expected. Nevertheless, the analysis should be sufficient to demonstrate whether GAs may be an appropriate method for solving this type of problem.

The partial differential equation that describes one-dimensional chemical transport in groundwater is (van Genuchten and Alves, 1982).

$$\frac{\partial}{\partial x}\left(\theta D \frac{\partial c}{\partial x} - \theta v c\right) - \frac{\partial}{\partial t}(\theta c + \rho_b s) = \mu_w \theta c + \mu_s \rho_b s - \gamma_w \theta - \gamma_s \rho_b \qquad (9)$$

where c is the solute concentration [ML^{-3}],

s is the sorbed solute concentration [MM^{-1}],

θ is the volumetric moisture content [$L^3 L^{-3}$],

D is the dispersion coefficient [$L^2 T^{-1}$],

v is the groundwater pore velocity [LT^{-1}],

ρ_b is the porous medium bulk density [ML^{-3}],

x is the distance [L],

t is the time [T],

μ_w is the liquid phase first-order decay constant [T^{-1}],

μ_s is the solid phase first-order decay constant [T^{-1}],

γ_w is the liquid phase zeroth order production constant [$ML^{-3}T^{-1}$],

and γ_s is the solid phase zeroth order production constant [T^{-1}].

The assumption of a linear sorption isotherm will be used in this research. Therefore,

$$s = kc \qquad (10)$$

where k is the partitioning coefficient.

The retardation (R) of the transported chemical is given by:

$$R = 1 + \frac{\rho k}{\theta} \qquad (11)$$

The rate coefficients μ and γ are defined as:

$$\mu = \mu_w + \frac{\mu_s \rho k}{\theta} \qquad (12)$$

$$\gamma = \gamma_w + \frac{\gamma_s \rho}{\theta} \qquad (13)$$

Substituting (10) – (13) into (9) yields:

$$D\frac{\partial^2 c}{\partial x^2} - v\frac{\partial c}{\partial x} - R\frac{\partial c}{\partial t} = \mu c - \gamma \qquad (14)$$

Disregarding first-order decay and production of the transported chemical and rearranging (14) yields the governing equation for the one-dimensional advective / dispersive transport model used in this research.

$$R\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} - v\frac{\partial c}{\partial x} \qquad (15)$$

van Genuchten and Alves (1982) provide the analytic solution to (15) assuming the following initial and boundary conditions.

The initial and boundary conditions are:

$$c(x,0) = C_i$$

$$c(0,t) = C_0 \qquad 0 < t < t_0$$

$$c(0,t) = 0 \qquad t > t_0$$

$$\frac{\partial c}{\partial x}(\infty,t) = 0$$

Note that the boundary condition at x = 0 specifies a pulse of chemical at concentration C_0 from time t = 0 to time t = t_0. The solution to (15) with the above initial and boundary conditions is,

$$c(x,t) = C_i + (C_0 - C_i)A(x,t) \qquad 0 < t < t_0 \qquad (16a)$$

$$c(x,t) = C_i + (C_0 - C_i)A(x,t) - C_0 A(x,t-t_0) \qquad t > t_0 \qquad (16b)$$

where,

$$A(x,t) = \frac{1}{2} erfc\left(\frac{Rx - vt}{2(DRt)^{1/2}}\right) + \frac{1}{2}\exp\left(\frac{vx}{D}\right) erfc\left(\frac{Rx + vt}{2(DRt)^{1/2}}\right)$$

2.3.2.2.2 Stochastic Modeling Approach

Field experiments have demonstrated that the heterogeneity in an aquifer can generally be represented stochastically by a lognormal distribution (Domenico and Schwartz, 1998). Enfield (2000) describes a modeling approach in which the flow field can be conceptualized as a bundle of flow tubes, each with different travel times. The tracer transport times in the flow field can be described as having a lognormal distribution with a mean travel time (*mtt*) and standard deviation (σ). The breakthrough curve can be modeled by defining the transport time of individual flow tubes as,

$$T_I = R * \exp \left(\text{normsinv} \left(I / N \right) * \sigma \right) \qquad (17)$$

where T_I is the travel time (normalized as pore volumes Q_I/n) for flow tube I,

$I = 1, 2, 3 \ldots (N\text{-}1)$,

Q_I is the Volumetric flow rate in tube I,

n is the porosity,

R is the retardation,

N is the number of flow tubes,

normsinv is the inverse of the standard normal cumulative distribution,

and σ is a unitless fitting parameter that describes the actual model standard deviation.

If the concentration is normalized as relative concentration ($C_r = C/C_0$, where C_0 equals the tracer concentration in the injected slug), the normalized travel times for individual flow tubes from (17) can be used to model the relative concentration at the extraction well. For time (t) less than the length of tracer injection time (t_0), the relative concentration at the extraction well can be modeled as a step function by summing the relative concentrations of the individual flow tubes. The relative concentration for an individual flow tube is $1/N$ if the tracer within that flow tube has reached the extraction well, i.e., the travel time (T_I) is less than the normalized time (t / mtt). If the tracer within the flow tube has not yet reached the extraction well, the relative concentration is zero. For t greater than t_0, the relative concentration at the extraction well can be modeled by including a negative step function beginning at t_0. The relative concentrations for the negative step function are determined in a similar manner to the step function except that travel times for individual flow tubes are compared to $(t - t_0) / mtt$. By the principle of

superposition, the relative concentration at the extraction well can be modeled by subtracting the negative step function from the step function (Figure 5).
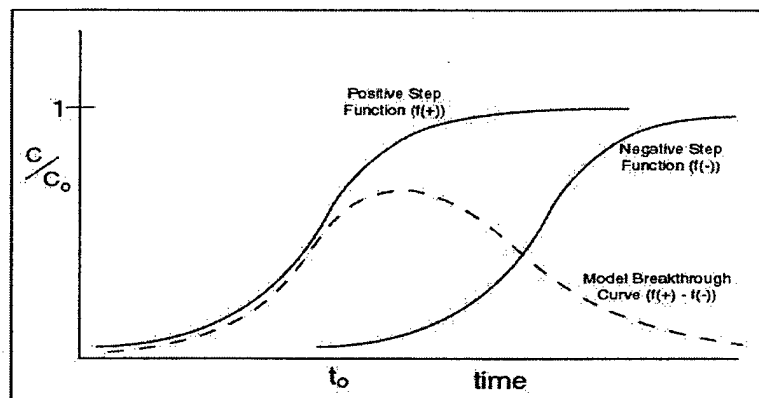
Figure 5: Step Functions and the Principle of Superposition

2.4 Genetic Algorithms

John Holland and his associates at the University of Michigan developed genetic algorithms (GAs), a mathematical optimization technique based on the principles of natural selection, in the 1960s. Early applications of GAs were in the realm of artificial intelligence and pattern recognition programs, but the flexibility of the GA makes it appropriate for a variety of optimization problems including a wide range of environmental and remediation modeling applications (Reeves, 1993).

2.4.1 GA Terminology and Methodology

Because GAs are based on the principles of natural selection, genetic terms are used. The process by which GAs determine an optimum solution set can most easily be described using analogies to genetics and reproduction. The user defines a *population* of decision variable sets, or *chromosomes* (also referred to as *strings* or *individuals*). The

27

decision variables that form the chromosomes are analogous to genes, or *alleles*. Alleles

are typically represented in binary code, although encoding can also use real or integer

values (Figure 6).
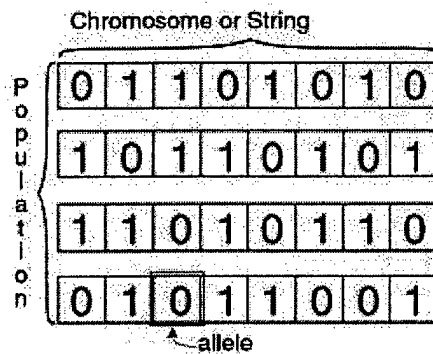
Figure 6: GA Terms

The user can define a value range and provide initial guesses for the variables to

be optimized. The GA encodes the initial values, along with additional randomly

selected values, into binary strings that form the initial population, the size of which is

defined by the user. The GA then applies the processes of *selection*, *crossover*, and

mutation to form subsequent populations. Selection occurs when the GA evaluates the

fitness of individuals within the population, and the fittest reproduce to form *offspring*.

The GA determines fitness through comparison to optimization criteria defined by the

user. For example, if minimizing an objective function in a model defines the

optimization criteria, the individuals within a population that return the lowest value for

that objective function are selected for reproduction. Reproduction occurs through

crossovers between randomly selected pairs that swap a portion of their gene string (the

crossover point is also randomly selected). The GA selects a percentage (determined by

the user) of the new population (composed of parents and offspring) for *mutation,* in

which the binary value of an allele changes. The processes of crossover and mutation

produce a new population comprised of the fittest individuals from the initial population

and their offspring. The GA evaluates the fitness of individuals within this new

population, and the process repeats until stop criteria are met (Figure 7).
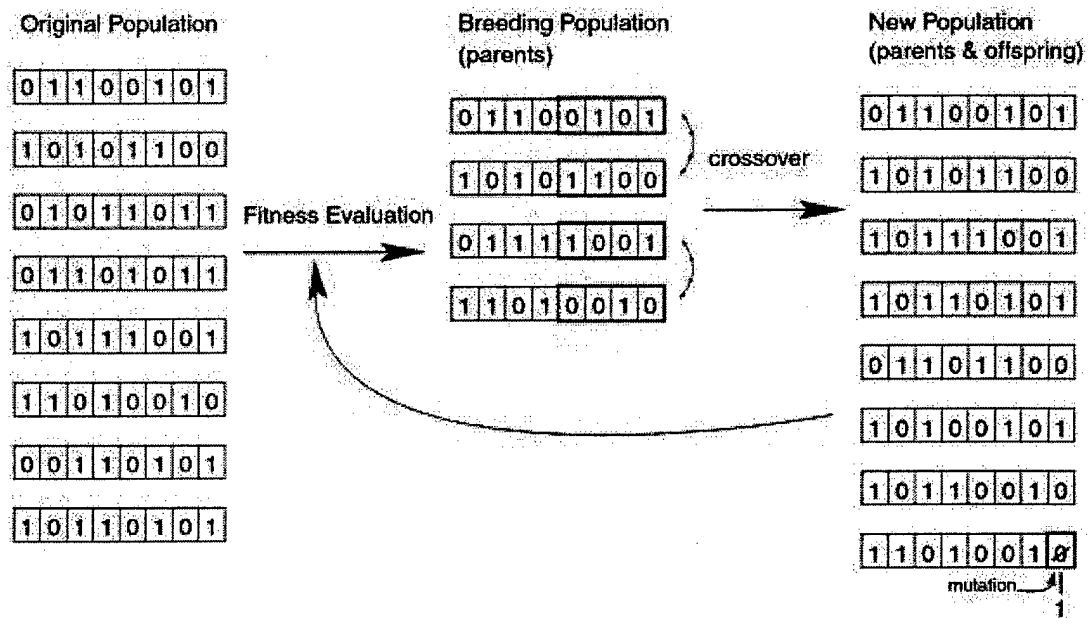
Figure 7: GA Process

2.4.2 Schema Theorem

The *schema theorem* provides the theoretical basis for the GA to evaluate fitness

of individuals within a population. A *schema* (plural *schemata*) defines a subset of the

population composed of similar individuals. For example, the chromosomes { 1 1 0 0 1 1

0 } and

29

{ 0 1 1 1 0 1 1 } are both members of the schema { * 1 * * * 1 * }. Both chromosomes will be members of several schemata that may contain one or both of the chromosomes. The categorization of individuals into schemata allows determination of an average fitness for each schema. This *intrinsic parallelism* allows the GA to evaluate the fitness of individuals within a population with fewer trials. Through crossover and mutation, the representation of a schema within a population will increase or decrease with relative fitness (Reeves, 1993).

2.4.3 Effects of GA Parameters on Performance

Knowledge of the decision variables that the GA will optimize can be used to provide initial guesses or ranges for parameter values that may enhance the performance of the GA. However, the possibility of early convergence to a solution that does not represent the global optimum exists. The required population size is related to the length of the binary code strings that comprise the population, but experience has indicated that populations of 30 individuals are adequate for most situations. Increasing the mutation rate decreases the probability of early convergence at a solution that does not represent the global optimum, but can significantly increase the time required for the GA to optimize the decision variables (Reeves, 1993).

2.4.4 Applications of GAs to Groundwater Remediation Problems

Several papers have been written about the application of GAs to groundwater remediation problems. Ritzel *et al.* (1994) developed a GA to handle multiple objective groundwater pollution containment problems and reported favorable results in designing a dual objective pumping system that maximized reliability and minimized cost. Garrett *et al.* (1999) applied a GA to a bioremediation problem for a TCE contaminated aquifer.

The GA was used to optimize a number of engineered parameters that affected the flow imposed by groundwater circulation wells, as well as biodegradation kinetics. Harrouni *et al.* (1996) combined GAs with the boundary element method for optimization of pumping well placement and groundwater parameter estimation. This research investigates the applicability of GAs to inverse modeling of PITT data.

2.5 Source of Data

The data used in this research was collected during field tests of the cleanup effectiveness of a variety of cosolvent / surfactant mixtures. The tests were conducted at Operable Unit 1 (OU1), Hill AFB, Utah (Rao *et al.*, 1997; Falta *et al.*, 1999, Jawitz *et al.*, 1998). OU1 is the site of several former disposal sites, including chemical disposal pits in which a variety of liquids including petroleum hydrocarbons such as jet fuel and chlorinated solvents were disposed (Falta *et al.*, 1999). The aquifer at the site is a shallow, unconfined aquifer approximately 6.1 m thick underlain by a thick clay unit that extends to depths greater than 90 m (Rao *et al.*, 1997). The tests involving cosolvent / surfactants floods were conducted within hydraulically isolated test cells constructed by driving interlocking sheet piling 2-3 m into the underlying clay layer (Figures 8 & 9). Pre-flood and post-flood PITT tests were used (in conjunction with other characterization methods) to estimate the spatial distribution of the NAPL and cleanup efficiency. Tests conducted at Cells 8 and F provide the data for this research.
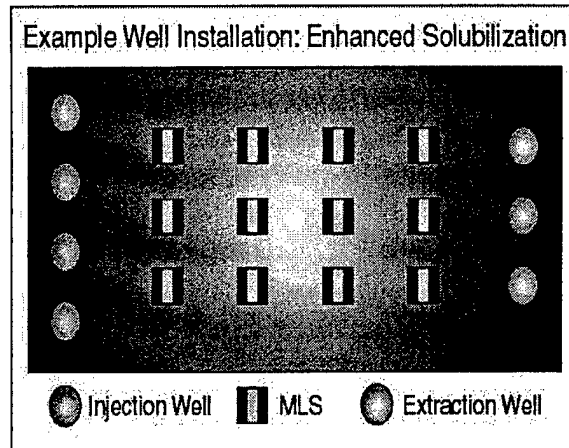
Figure 8: Plan View of Test Cell Configuration
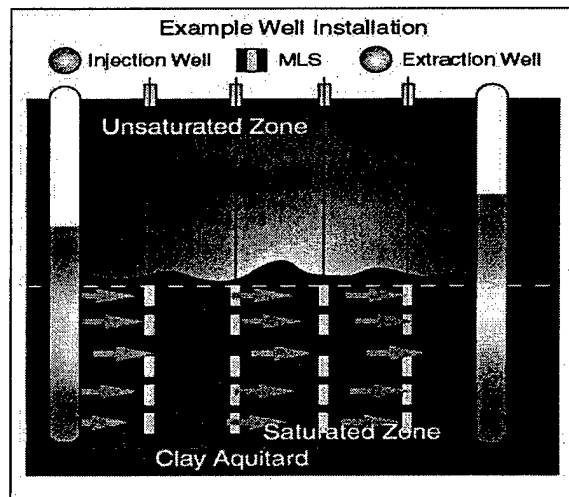
Figure 9: Elevation View of Test Cell Configuration

2.5.1 Cell 8

Tests conducted by researchers from the University of Florida, in conjunction with the

32

US EPA and USAF (Jawitz *et al.*, 1998), examined the effectiveness of a Windsor Type I

surfactant/alcohol (surfactant/co-surfactant) mixture as a single-phase micro-emulsion

(SPME) in a hydraulically isolated test cell measuring 2.8m x 4.6 m. The composition of

the NAPL present at this site is shown in Table 1. Pre- and post-flood NAPL saturation

was determined using soil samples and the PITT technique. Results indicated removal of

90 – 95% of the most prevalent NAPL components.

Target Analyte	Mass Fraction, (g/100g NAPL)
p-Xylene	.144
1,2,4-Trimethylbenzene	.4388
n-Decane	.477
n-Undecane	1.573
n-dodecane	.698
n-tridecane	.285

Table 1: Target Analyte Mass Fractions in the NAPL (Jawitz *et al.*, 1998)

2.5.2 Cell F

Tests conducted by researchers from the University of Florida, in conjunction

with the US EPA and USAF (Rao *et al.*, 1997), examined the efficiency of a cosolvent

solubilization flood, consisting of water and two water-miscible alcohols, as a

remediation technique for a NAPL-contaminated aquifer. The tests were conducted

within a hydraulically isolated test cell measuring approximately 4.3 m x 3.5 m. Pre-

flood characterization of the test cell was conducted using soil cores and groundwater

samples, and the volume and distribution of the NAPL was estimated with the PITT

technique. The composition of the NAPL contaminant present at the site is shown in

Table 2. Following the cosolvent flood, soil cores and groundwater samples were

collected to determine the concentrations of NAPL constituents remaining. Results indicated nearly a 90 – 99% removal in the upper 1 m zone, and removal efficiencies dropped to 70 – 80% in the bottom 0.5 m above the confining clay layer. Data from the pre-flood PITTs were reported by Annable *et al.* (1997).

Target Analyte	Mass Fraction, x 10^3
1,1,1-Trichloroethane	0.016
p-,m-Xylene	0.17
1,2-Dichlorobenzene	6.1
Toluene	0.074
1,3,5-Trimethylbenzene	0.83
n-Decane	5.2
n-Undecane	16
Napthalene	0.11

Table 2: Target Analyte Concentrations in the NAPL (Rao *et al.*, 1997)

Chapter 3. Experimental Method

3.1 Introduction

The application of GAs as a "best fit" parameter determination method to help interpret PITT results was evaluated using two modeling approaches. The first approach applied a GA to estimate flow parameters (v, D, and R) for the analytic solution to the one-dimensional advective / dispersive equation (16). The second approach used a GA to estimate fitting parameters for a stochastic model based on the assumption of a lognormal distribution of travel times for the breakthrough curve (17). Parameters were optimized to fit model simulations to breakthrough curve data for partitioning and conservative tracers obtained from PITTs at two test cells at Hill AFB, Utah. The results obtained from the GA model were used to estimate the NAPL saturation and cleanup efficiency of cosolvent / surfactant floods at those test cells. The NAPL saturation and cleanup efficiency estimates from the GA models were compared to estimates obtained directly from the experimental data using the method of moments.

3.2 GA Models

3.2.1 Evolutionary Solver

The GA used in this model is an optimization tool included with Premium Solver from Frontline Systems (Incline Village, NV), an add-in for Microsoft Excel. For this research, the educational version of Premium Solver was used. The educational version differs from the full version in that it limits the number of variables to be optimized to 250 (compared to 5000 for the full version), but the capabilities of the educational version are sufficient for this research. Figure 10 shows the Premium Solver Parameters window. Note that the Standard Evolutionary optimization method is selected. At this

interface, the user defines the objective function, optimization constraint, variables, and variable constraints. The sample optimization problem in Figure 10 shows that the objective function is defined in cell $N$9. The GA is directed to find the minimum value for the objective function by changing the decision variables (v, D, n) located in cells $D$21:$D$31. Upper and lower bounds are defined for the decision variables, and an integer constraint is placed on "n".
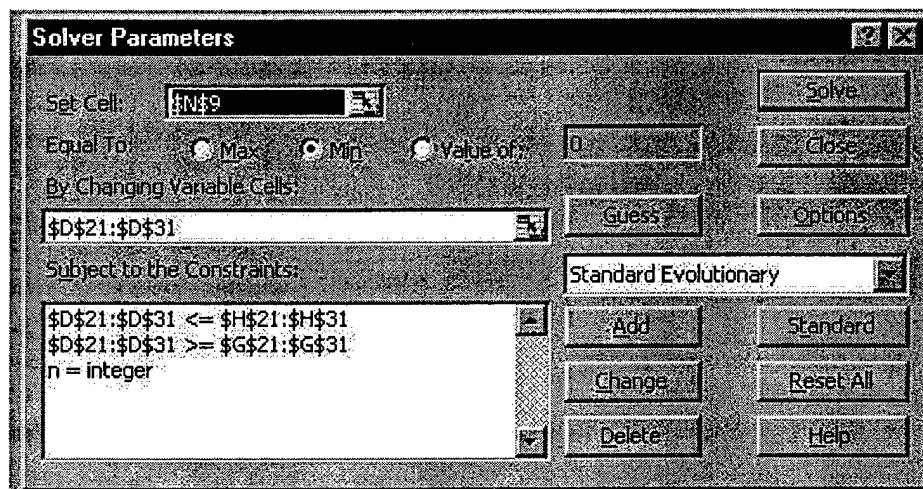
Figure 10: Premium Solver for Excel User Interface

The Solver Options window (Figure 11) allows the user to define how the GA will function. The Max Time and Iterations limits become a factor if the GA does not converge on an optimum solution. The limits stop the program in a reasonable amount of time and allow the user to determine whether to continue or adjust other options to improve performance. Precision helps define the length of the encoded strings and can affect the time required for the GA to perform its search and optimization routine. Because of this, the precision required should be carefully evaluated and minimized.

Convergence defines the relative improvement in fitness that must be achieved for the GA to continue. The effects of population size and mutation rate on GA performance were discussed in Section 2.4.3. A larger population and higher mutation rate increase the chances of finding a global rather than a local minimum for the objective function, but can significantly impact the length of time required for the GA to converge on a solution.
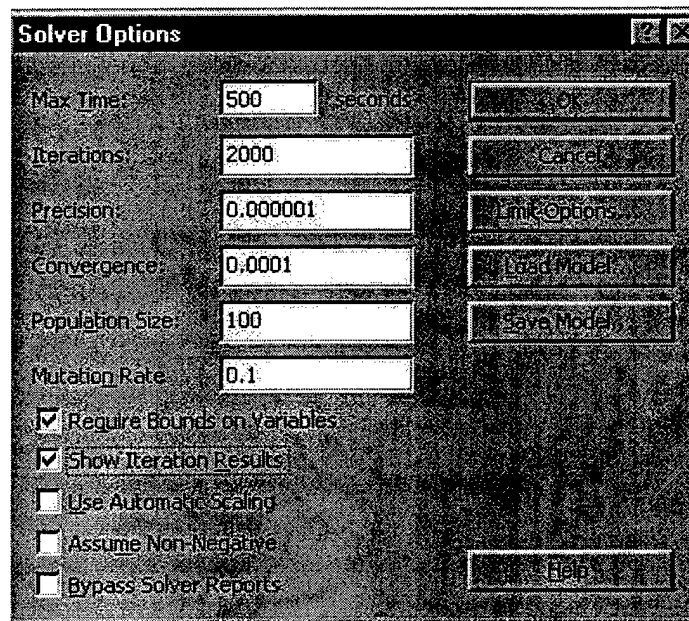
Figure 11: Solver Options Window

3.2.2 Objective Function

The objective function that was minimized in the model was the sum of the Chi Square function for each data point.

$$X^2 = \sum_{i=1}^{k} \frac{(\exp_i - \text{mod}_i)^2}{\exp_i} \qquad (18)$$

where $\exp_i$ is the experimental (measured) tracer concentration at time t_i,

mod_i is the model tracer concentration at time t_i,

and   n is the number of data points.

The Chi Square function was chosen because it accounts for the scaled difference between the model and PITT data.

3.2.3 Analytic Model

The analytic model used discrete values for hydraulic and chemical parameters for the analytic solution to the one-dimensional advective / dispersive equation described in Section 2.3.2.2.1. Several simplifying assumptions, discussed in Section 2.3.2.2.1, were made in this analysis to minimize the complexity of the analytic model and the number of parameter values that the GA would be required to optimize. The concentration of the injected tracer solution (C_0), the tracer injection time (t_0), and the distance between the injection and extraction wells (x) were constants in the analytic model. It should be noted that assuming that C_0 is constant would affect the accuracy of the model since mass loss to other extraction wells or through degradation is not taken into consideration. As a result, the model may tend to overestimate retardation and NAPL saturation. Heterogeneity in the aquifer, and the corresponding variation in flow parameters, were accounted for by applying the deterministic parameter values to a discrete number of layers of equal thickness.

38

3.2.3.1 Conservative Tracer Model

The breakthrough curves for the conservative tracer (both pre- and post-remediation) were used to estimate the hydraulic parameters (groundwater pore velocity and dispersion coefficient) and optimize the number of layers using Evolutionary Solver (Figure 12).
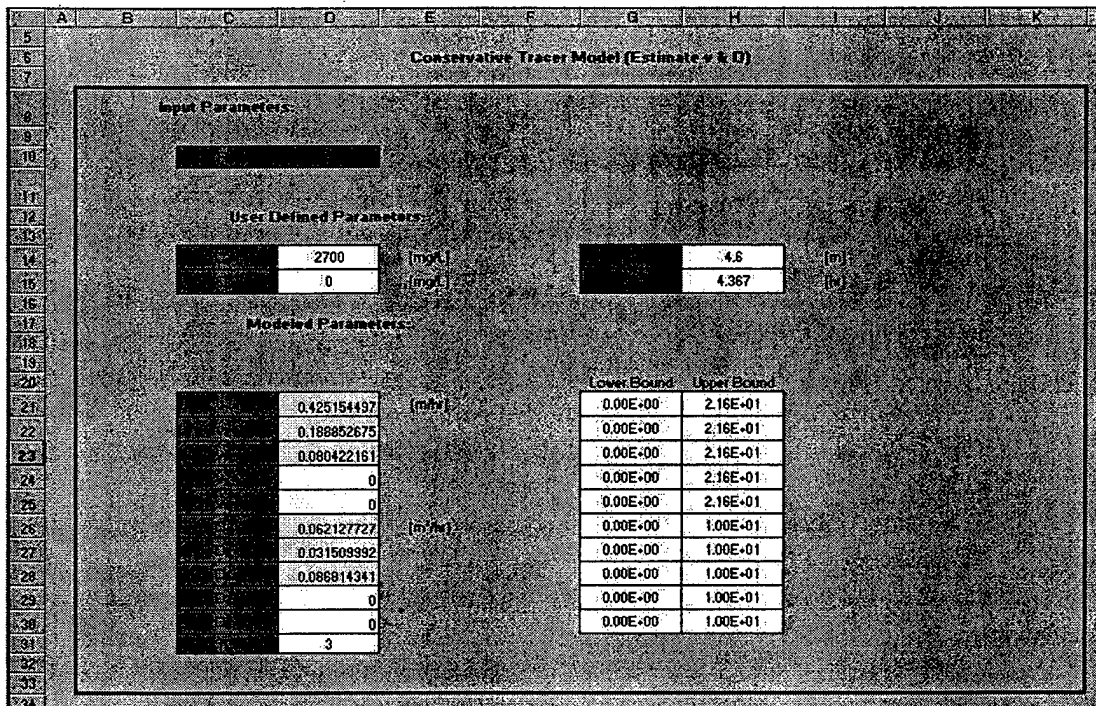
Figure 12: Conservative Tracer (Analytic Model)

3.2.3.1.1 Optimizing the Number of Layers

Because Genetic Algorithms utilize a random search to assign values to alleles, linking constraints are necessary to drive the GA towards optimizing both parameter values and number of layers. Such constraints are necessary because the GA assigns values for the flow parameters (v_l and D_l) in all five layers even though the GA may not

39

have selected all five layers (i.e. $n < 5$). For example, if $n = 4$, the GA has assigned

values to v_5 and D_5 even though those layers are not included in the analytic calculations.

As a result, the values for v_5 and D_5 may not provide a good fit of the model to the PITT

data, but they do not affect the fitness of the individual because

they are not included in the analytic calculations. However, if layer five is selected in a

subsequent generation, the values that have already been assigned to the layer parameters

hurt the fitness of the individual, preventing it from being selected for crossover. Over

multiple generations, this creates a bias towards optimizing the objective function with

one layer. To drive the GA to investigate potential solutions with more than one layer, a

penalty is added to the objective function for each layer that has not been selected but has

parameter values assigned. To allow this linking constraint to be incorporated into the

model, layer variables (Y1, Y2, …, Y5) were created, and logic constraints in the form of

penalties to the objective function were added to ensure layers were selected sequentially

(Figure 13).

	W	X	Y	Z	AA	AB	AC	AD
9								
10			Binary Variables for Problem					
11	Y1	Y2	Y3	Y4	Y5	Enforce Logical Conditions between Selected Layers		
12	1	1	0	0	0	Logic Constraint	Less Than	Penalty
13	-1	1					0	0
14		-1	1				0	0
15			-1	1			0	0
16				-1	1		0	0
17								0
18								
19								
20	Enforce Link Between Layer Variables and Number of Layers (n)							
21	Linking Constraints		Less Than	Penalty				
22	Layer 1		0	0				
23	Layer 2		0	0				
24	Layer 3		0	1000				
25	Layer 4		0	1000				
26	Layer 5		0	1000				
27			Total Penalty:	3000				
28								

Figure 13: Logic and Linking Constraints

Increasing the number of layers improves the fit of the model breakthrough curves to the experimental data, but there is a corresponding increase in model complexity. To achieve the goal of optimizing the fit of the model breakthrough curve while minimizing model complexity, a penalty was added to the objective function as the number of layers increased. Several model runs using this approach indicated a bias in the model for the GA to optimize the parameters using five layers. This bias stemmed from contradictory requirements for penalty values assigned to linking constraints and is discussed in detail in Section 4.2.1. As a result, individual model runs were performed for one to five layers. The results obtained from these runs were analyzed to determine the optimum number of layers that balanced goodness of fit with model complexity. The resulting number of layers was used in subsequent runs.

3.2.3.2 Partitioning Tracer Model

To estimate the retardation of the partitioning tracer in pre- and post-flood PITTs, the linear pore velocity and dispersion coefficient values estimated from the conservative tracer model were held constant and the GA was applied to optimize the model by varying the retardation value in each of the layers. Figure 14 shows the layout for the partitioning tracer model.
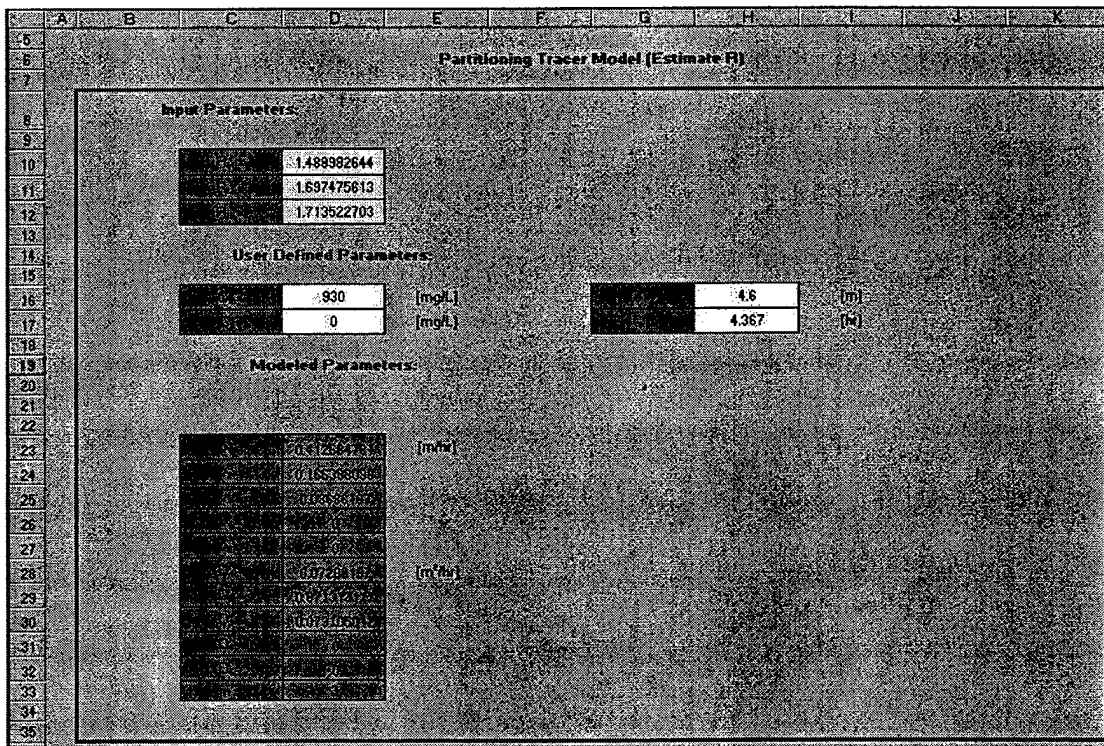
Figure 14: Partitioning Tracer Model (Analytic Model)

The NAPL saturation in each layer may be calculated using a modified form of equation (2) that gives NAPL saturation as a function of retardation.

$$(S_n)_i = \frac{R-1}{K_{nw} + R - 1} \tag{19}$$

The assumption that the model layers have equal thickness is implicit in the model, so the average NAPL saturation was calculated by;

$$(S_n)_{ave} = \frac{\sum_{i=1}^{n} (S_n)_i}{n} \tag{20}$$

where n is equal to the number of layers.

3.2.4 Stochastic Model

The ability of the GA to optimize parameter values for the stochastic modeling approach described by Enfield (2000) was evaluated using an Excel model and Evolutionary Solver. C_0 and t_0 were input to the model as constants; the problems associated with the assumption of a constant C_0 are discussed in Section 3.2.3. The value of stochastic function was determined for 100 flow tubes in each layer. Arrays were constructed that assigned a value of zero or one to each flow tube as a function of time based on the following criteria:

If the stochastic function is defined as:

f(I) = R * EXP (NORMSINV (I/N) * σ)

43

where EXP is the exponential function in Excel,

NORMSINV is the function in Excel that returns the inverse of the

standard normal cumulative distribution,

I is the number identifying the individual flow tube,

N is the number of flow tubes,

R is the retardation of the tracer,

and σ is the standard deviation of the stochastic function that describes the

breakthrough curve.

At a given time, t, the cell values within the array for the positive step function are

determined by:

for I = 1 . . . 99

if (f(I) < t/mtt)

cell value = 1

else

cell value = 0

At a given time, t, the cell values within the array for the negative step function

are determined by:

for I = 1 . . 99

if (f(I) < (t − t_0)/mtt)

cell value = 1

else

cell value = 0

44

where mtt is the mean travel time for the tracer

and t_0 is the injection time for the tracer.

The normalized model concentrations at a given time are determined by summing the cell values associated with each flow tube and dividing by the number of flow tubes.

3.2.4.1 Conservative Tracer Model

The breakthrough curves for the conservative tracer (both pre- and post-remediation) were used to estimate the standard deviation of the stochastic function and the mean travel time for the tracer. Parameter values were estimated for each layer. Figure 15 shows the layout for the conservative tracer model with three layers.
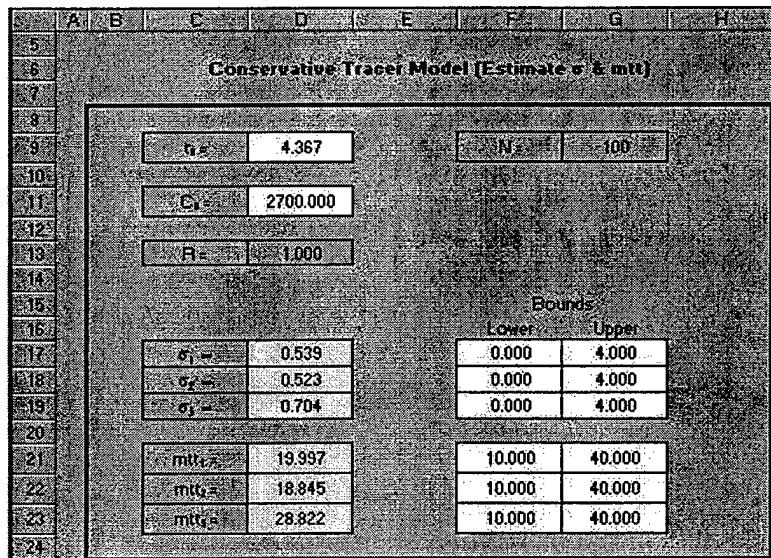
Figure 15: Conservative Tracer Model (Stochastic Model)

3.2.4.2 Partitioning Tracer Model

As in the analytic model, the parameter estimates from the pre- and post-flood stochastic conservative tracer models were held constant in the partitioning tracer model and the GA was applied to optimize the model by varying the retardation values in each of the layers. NAPL saturation was estimated using the method described in Section 3.2.2.2. Figure 16 shows the layout of the stochastic partitioning tracer model.
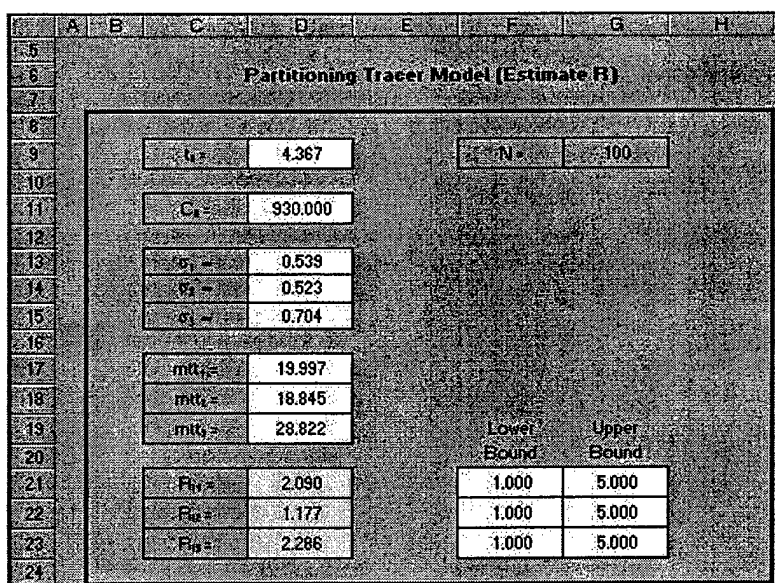
Figure 16: Partitioning Tracer Model (Stochastic Model)

3.3 Inverse Modeling of Experimental Data

The method of moments, described in Section 2.3.2.1, was used to estimate the mean travel time from the experimental data and calculate NAPL saturation and cleanup efficiency. The data were extrapolated to account for retardation in the tail. The inverse modeling results were compared to the results obtained from the analytic and stochastic models.

46

Chapter 4. Analysis

4.1 Overview

The use of GAs as a parameter optimization method for PITT interpretation was evaluated by determining the goodness of fit of model breakthrough curves to PITT experimental data. The utility of the models as a design and decision making tool was evaluated by comparing the estimates for NAPL saturation and cleanup efficiency obtained from the models to those calculated using the method of moments.

4.2 Analytic Model

4.2.1 Optimizing the Number of Layers

As discussed in Chapter 3, initial results using the analytic model indicated that conflicting constraints prevented the GA from returning an optimum number of layers for the model that met the desired criteria. Large linking constraint penalties to the objective function associated with assigning parameter values to non-selected layers were necessary to drive the GA towards examining solutions with multiple layers. On the other hand, it was also necessary to incorporate penalties for extra layers, in order to minimize the number of fitting parameters. However, these penalties would be necessarily small when compared to the linking constraint penalties. As a result, they had no impact on the model, and the GA continued to demonstrate a bias towards maximizing the number of layers.

Individual model runs were conducted using the pre-flood conservative tracer data from Cell 8 to empirically determine the number of layers that would be used in subsequent evaluations. Results indicated significant improvement when the number of layers was increased to three, with only minor improvement in model fit with further

increases in the number of layers (Figure 17). Because the geometry of the breakthrough curves was similar for all data sets, subsequent model runs were based on the assumption that three layers would allow a reasonable fit of the models to the experimental data. This decision was based on the premise that increasing the number of layers significantly impacted the amount of time required to optimize the model parameters and, in principle, less fitting parameters are preferable.
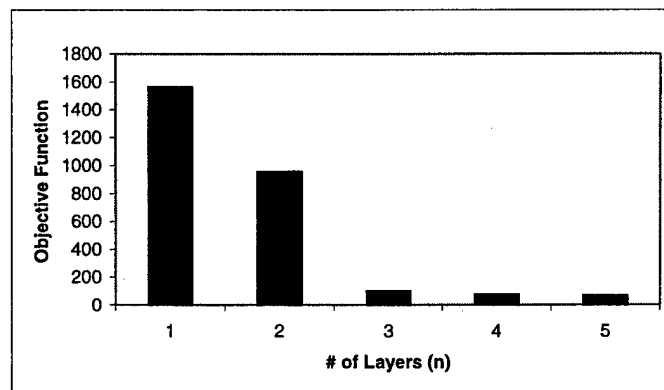
Figure 17: Value of Minimized Objective Function with Increasing "n"

4.2.2 Cell 8

The breakthrough curves representing the analytic solution to steady state one-dimensional advective dispersive equation (hereafter referred to as *model breakthrough curves*) are represented as a solid line along with the experimental data points in Figure 18. The results demonstrated that the analytic model breakthrough curve provided a good fit to the experimental data. The extraction well data indicated a bimodal peak in concentration for the pre-flood conservative tracer breakthrough curve and the pre- and post-flood partitioning tracer breakthrough curves. Modeling the breakthrough curves

using superposition of multiple layers with discrete parameter values accounted for this

characteristic of the extraction well data.
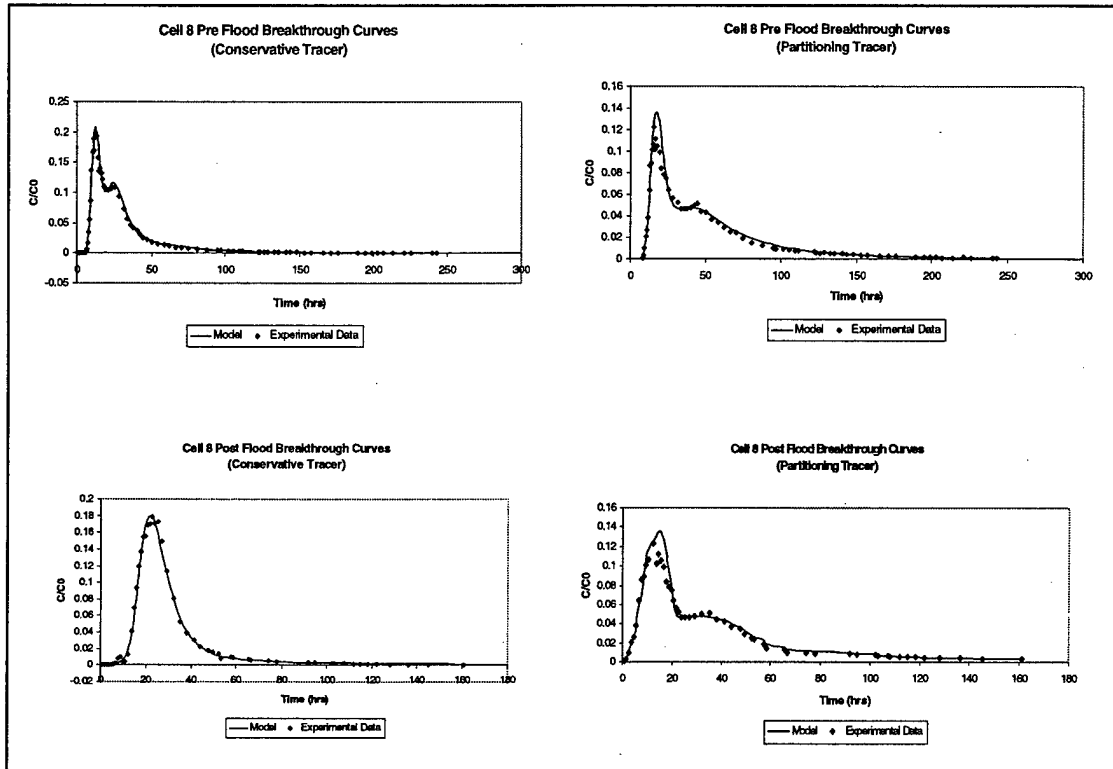
Figure 18: Analytic Model Breakthrough Curves and Experimental Data (Cell 8)

Estimates for flow parameters developed by the model (Table 3) were within the

expected range for a sand and gravel aquifer. Jawitz *et al.* (1998) reported an average

hydraulic conductivity of 0.36 m/hr across the test cell, and an effective porosity of 0.14.

Jawitz *et al.* did not report the hydraulic gradient for Cell 8. However, the average linear

pore velocity estimated by the model (0.23 m/hr) would be consistent with the results

49

reported by Jawitz *et al.* for a hydraulic gradient of approximately 0.1m/m, a reasonable value.

Because removal of NAPL mass would be expected to increase the relative permeability of the aquifer, an increase in velocity estimates was expected for the post-flood models. As seen in Table 3, this was not found. However, the model results did show a general correlation between layer velocity and NAPL saturation. Lower velocities were associated with higher R values as would be expected if the presence of NAPL decreases relative permeability.

A longitudinal dispersivity of approximately 0.1 m would be expected based on the scale of the test cell experiment (Domenico and Schwartz, 1998). The dispersion values for Layer 1 and Layer 2 in the pre-flood model and for Layer 1 and Layer 3 in the post-flood model are consistent with this expectation. The GA returned a value for the dispersion coefficient for Layer 2 in the post-flood model, but this value was not included in the calculations (since $v_2 = 0$) and did not affect the model fit to the data.

Cell ID:	8				
EW	1				
Model Type	Analytic				
Opt Method	GA				
			Parameter Estimates		
Test Type	Tracer Type	Layer ID	v (m/hr)	D (m^2/hr)	R
Pre Flood	Conservative	1	0.425	0.062	
		2	0.189	0.032	
		3	0.080	0.087	
Pre Flood	Partitioning	1			1.489
		2			1.697
		3			1.714
Post Flood	Conservative	1	0.169	0.045	
		2	0.000	0.196	
		3	0.224	0.034	
Post Flood	Partitioning	1			1.123
		2			1.000
		3			1.071

Table 3: Analytic Model Parameter Estimates (Cell 8)

Parameters fitted to the partitioning tracer data returned retardation (R) values

indicative of NAPL mass removal (Table 3). The model pre- and post flood

breakthrough curves are shown in Figure 19. Note that the partitioning tracer breaks

through before the conservative tracer in the post-flood breakthrough curves, but due to

tailing has a retardation factor greater than one. These characteristic of the breakthrough

curve for the partitioning tracer (early breakthrough and tailing) may be indicative of rate

limited sorption (Brusseau and Rao, 1989).
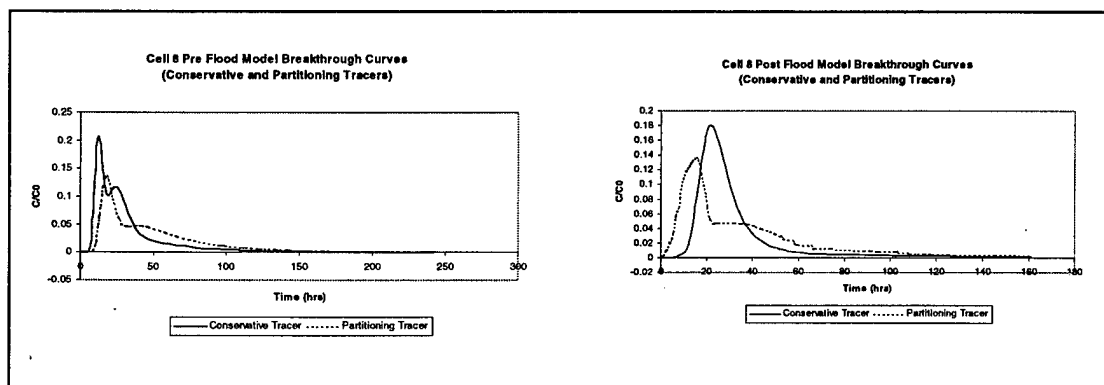
Figure 19: Analytic Model Pre- and Post-Flood Breakthrough Curves (Cell 8)

Table 4 shows the pre- and post-flood NAPL saturation, as well as the cleanup

efficiency, estimated by the model. Because v for layer 2 was optimized at zero, the flow

field has essentially been modeled with two layers in the post-flood model. As a result,

layer 2 was not included when calculating the average post-flood saturation.

Because Jawitz et al. (1998) reported only pre- and post-flood concentrations for

the target NAPL constituents in their study, their results do not provide a basis for

comparing the NAPL saturation estimated by the model. Additionally, the model

estimates for NAPL saturation and cleanup efficiency used data from only one extraction

well. Since the estimates from one extraction well should not be considered

representative of the entire test cell, a direct comparison to the results reported by Jawitz

et al. (1998) is not possible. However, a comparison can indicate whether the model is

providing a reasonable representation of the experimental data. Jawitz *et al.* (1998)

reported a cleanup efficiency of 72 %. It is likely that their results more accurately reflect

actual cleanup efficiency since, beyond the reasons discussed above, the estimates were

obtained using data from both the PITTs and core samples.

Cell ID:	8
EW	1
Model Type	Analytic
Opt Method	GA

Partitioning Tracer	2,2-dimethyl-3-pentanol
K_{nw}	7.42

Test Type	Layer ID	S_n	Average S_n
Pre Flood	1	0.0618	0.0785
	2	0.0859	
	3	0.0877	
Post Flood	1	0.0163	0.0129
	2	0.0130	
	3	0.0095	
	Cleanup Efficiency		83.57%

Table 4: Analytic Model Estimate of NAPL Saturation and

Cleanup Efficiency (Cell 8)

4.2.3 Cell F

Model breakthrough curves for Cell F demonstrated a reasonable fit to the experimental data (Figure 20). The bimodal peak of the breakthrough curves observed in the models of the Cell 8 data was also observed in the post-flood models for Cell F, and was represented by the model.
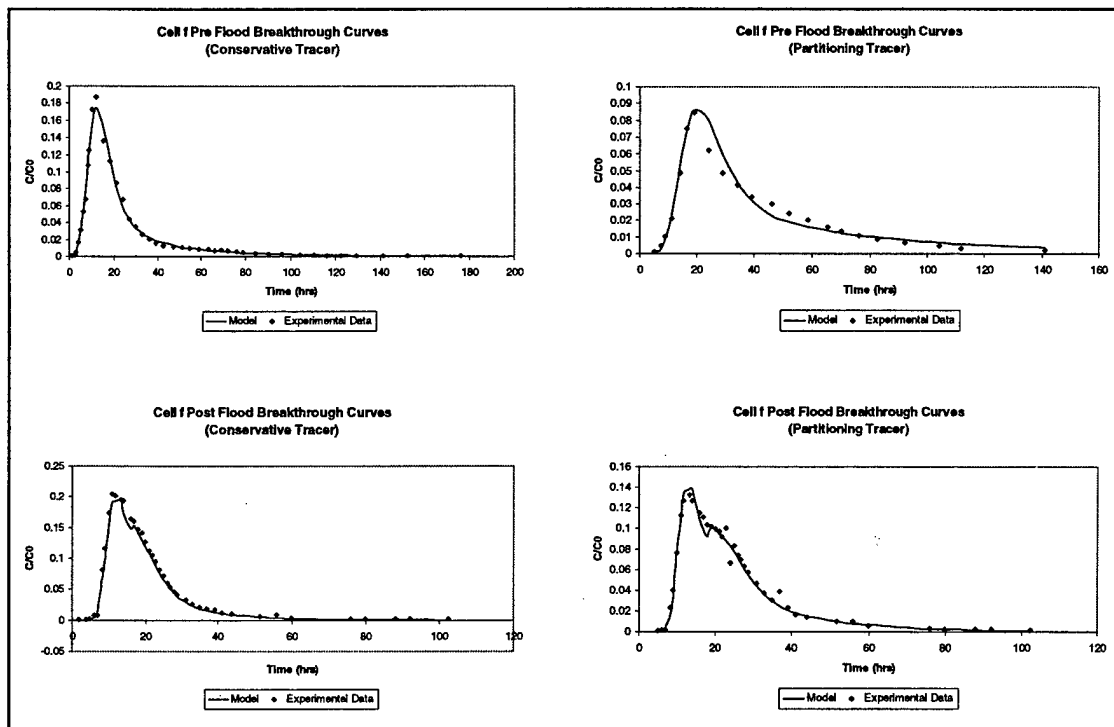
Figure 20: Analytic Model Breakthrough Curves and Experimental Data (Cell F)

Estimates for flow parameters developed by the model (Table 5) were within the expected range for a sand and gravel aquifer. The average linear pore velocity estimated by the model was 0.22 m/hr. The hydraulic conductivity reported by Rao *et al.* (1997) was 0.72 m/hr, and the effective porosity was reported as 0.20. The linear pore velocity

estimates from the model would be consistent with those reported by Rao *et al.* (1997) if the hydraulic gradient is 0.06, a reasonable value (though not reported).

The model estimates for the post-flood velocities reflected the expectation that increased velocities may result from the increase in effective permeability of the aquifer following NAPL mass removal. The model results also showed the general inverse relationship between velocity and NAPL saturation that was seen in the results from Cell 8.

Cell ID:	f
EW	3
Model Type	Analytic
Opt Method	GA

Test Type	Tracer Type	Layer ID	Parameter Estimates		
			v (m/hr)	D (m²/hr)	R
Pre Flood	Conservative	1	0.317	0.093	
		2	0.088	0.197	
		3	0.249	0.345	
Pre Flood	Partitioning	1			1.636
		2			2.970
		3			2.613
Post Flood	Conservative	1	0.417	0.052	
		2	0.169	0.117	
		3	0.239	0.038	
Post Flood	Partitioning	1			1.136
		2			1.477
		3			1.215

Table 5: Analytic Model Parameter Estimates (Cell F)

Optimization of the pre- and post-flood analytic model returned retardation (R) values indicative of NAPL mass removal (Table 5). Figure 21 shows the model pre- and post-flood breakthrough curves for the conservative and partitioning tracers.
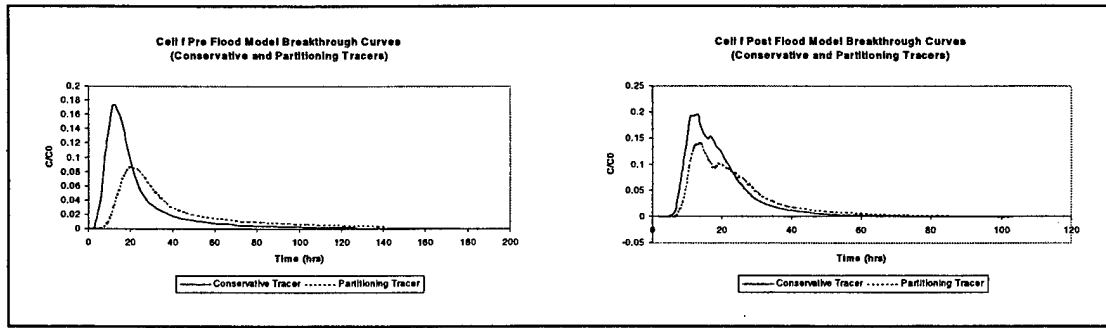
Figure 21: Analytic Model Pre- and Post-Flood Breakthrough Curves (Cell F)

Table 6 shows the pre- and post-flood NAPL saturation, as well as the cleanup

efficiency, estimated by the model. The cleanup efficiency of 77.19 % estimated by the

model is comparable to the cleanup efficiency of 82.61 % estimated from PITT results

and reported by Rao *et al.* (1997). Reasons for the differences between the model and

reported data were discussed in Section 4.2.2.

Cell ID:	f
EW	3
Model Type	Analytic
Opt Method	GA

Partitioning Tracer	2,2-dimethyl-3-pentanol
K_{nw}	7.42

Test Type	Layer ID	S_n	Average S_n
Pre Flood	1	0.0789	0.1558
	2	0.2098	
	3	0.1786	
Post Flood	1	0.0180	0.0355
	2	0.0604	
	3	0.0282	
	Cleanup Efficiency		77.19%

Table 6: Analytic Model Estimate of NAPL Saturation and

Cleanup Efficiency (Cell F)

55

4.3 Stochastic Model

4.3.1 Cell 8

Figure 22 shows the stochastic model simulations compared to experimental breakthrough data from Cell 8. The models of the pre-flood breakthrough curves provided a reasonable fit to the experimental data, although the model tended to underestimate the peak concentrations. Model runs performed by Enfield (2000) in which the number of flow tubes were varied indicated that increasing the number of flow tubes enables the model to account for peak concentrations by "smoothing" the curve. The model in this research was limited to using a total of 300 flow tubes (three layers, each with N=100 flow tubes) because, due to the size of the arrays in Excel, the GA was unable to optimize the model fitting parameters for N>100.

Figure 22: Stochastic Model Breakthrough Curves and Experimental Data (Cell 8)

Table 7 shows the stochastic parameter estimates for Cell 8. The model demonstrated the same inverse relationship between layer velocity (reflected in the layer mean travel time (mtt)) and NAPL saturation (indicated by R) as was seen when the analytic model was applied to Cell 8.

Cell ID:	8
EW	1
Model Type	Stochastic
Opt Method	GA

			Parameter Estimates		
Test Type	Tracer Type	Layer ID	σ	mtt (hrs)	R
Pre Flood	Conservative	1	0.539	19.997	
		2	0.523	18.845	
		3	0.704	28.822	
Pre Flood	Partitioning	1			2.090
		2			1.177
		3			2.286
Post Flood	Conservative	1	0.673	28.224	
		2	0.413	35.699	
		3	0.688	24.617	
Post Flood	Partitioning	1			1.310
		2			1.000
		3			1.441

Table 7: Stochastic Model Parameter Estimates (Cell 8)

Figure 23 shows the separation of the conservative and partitioning breakthrough curve from the pre- and post-flood stochastic models for Cell 8.
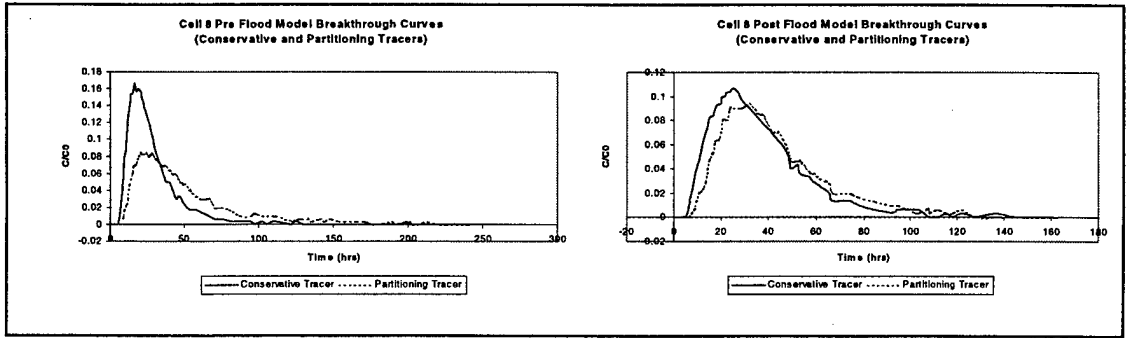
Figure 23: Stochastic Model Pre- and Post- Flood Breakthrough Curves (Cell 8)

Using moments calculated from the stochastic model for Cell 8, a cleanup

efficiency of approximately 68 % was determined (Table 8). The results obtained in this

model are compared with the stochastic model and the results from Jawitz *et al.* (1998) in

Section 4.4.

Cell ID:	8		
EW	1		
Model Type	Stochastic		
Opt Method	GA		

Partitioning Tracer	2,2-dimethyl-3-pentanol	
K_{nw}	7.42	

Test Type	Layer ID	S_n	Average S_n
Pre Flood	1	0.1280	0.0997
	2	0.0233	
	3	0.1477	
Post Flood	1	0.0401	0.0321
	2	0.0000	
	3	0.0561	
	Cleanup Efficiency		67.79%

Table 8: Stochastic Model Estimate of NAPL Saturation and

Cleanup Efficiency (Cell 8)

4.3.2 Cell F

The stochastic model provided a better fit to the experimental data for Cell F (Figure 24) than for Cell 8. However, for the reasons discussed in Section 4.3.1, the model tends to underestimate the peak tracer concentrations.
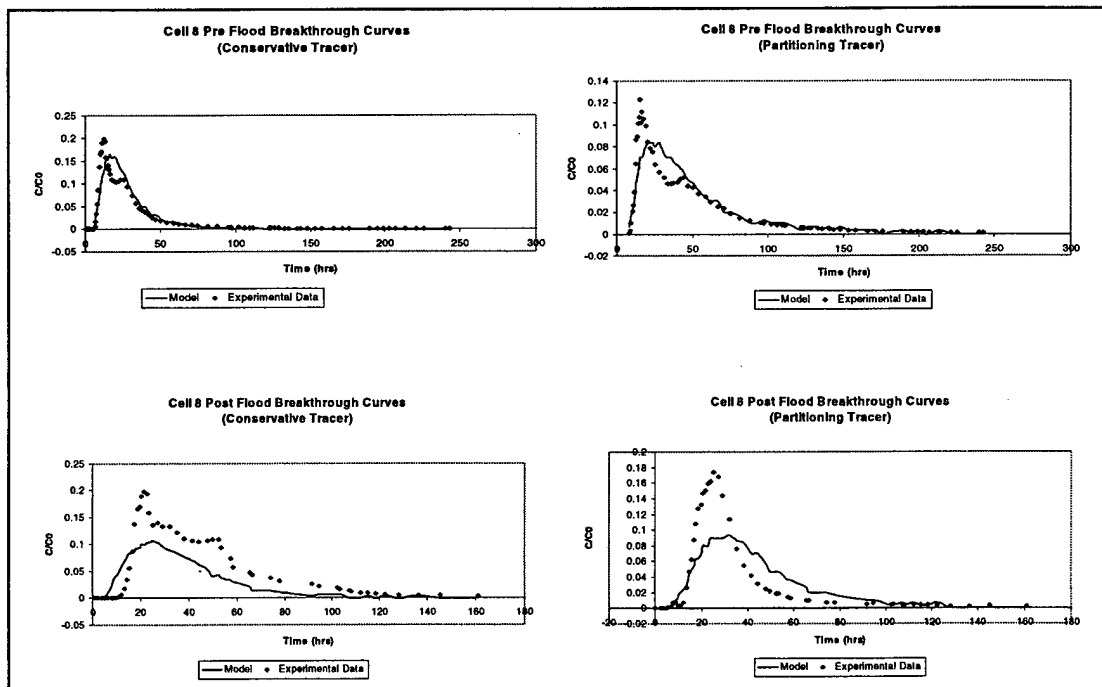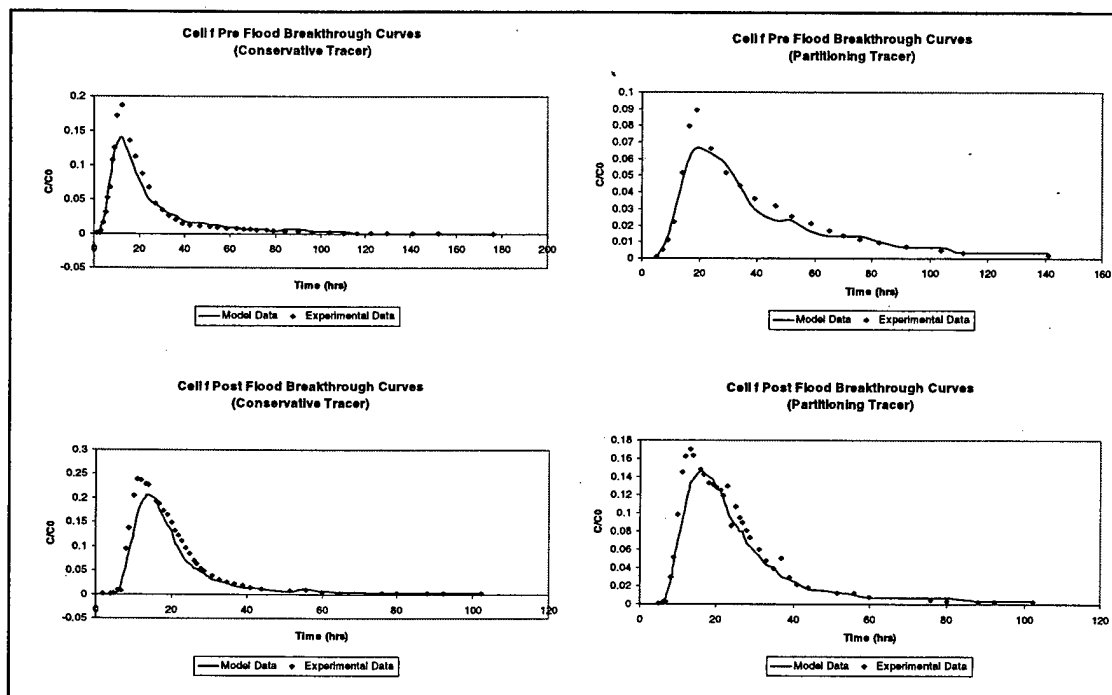

Figure 24: Stochastic Model Breakthrough Curves and Experimental Data (Cell F)

Table 9 shows the stochastic parameter estimates for Cell F. The correlation between mean travel time and NAPL saturation that was seen in the results for Cell 8 are not evident in the results from Cell F. The pre- and post-flood mean travel times for the conservative tracer demonstrate an interesting discrepancy from the expected performance of the model, with a sharp increase for the post-flood mean travel time in layer 1. This discrepancy may represent an artifact of the three layer model. Because the

model is not a true representation of field conditions, but an artificial representation to fit

experimental extraction well data, the best fit parameters may contradict expected

performance.

Cell ID:	f				
EW	3				
Model Type	Stochastic				
Opt Method	GA				
			Parameter Estimates		
Test Type	Tracer Type	Layer ID	σ	mtt (hrs)	R
Pre Flood	Conservative	1	0.430	11.989	
		2	1.031	31.242	
		3	1.110	26.225	
Pre Flood	Partitioning	1			1.931
		2			1.842
		3			3.718
Post Flood	Conservative	1	0.818	31.971	
		2	0.378	13.110	
		3	0.365	14.386	
Post Flood	Partitioning	1			1.427
		2			1.051
		3			1.483

Table 9: Stochastic Model Parameter Estimates (Cell F)

Figure 25 shows the separation of the model breakthrough curves for the

conservative and partitioning tracers for Cell F (pre- and post-flood).

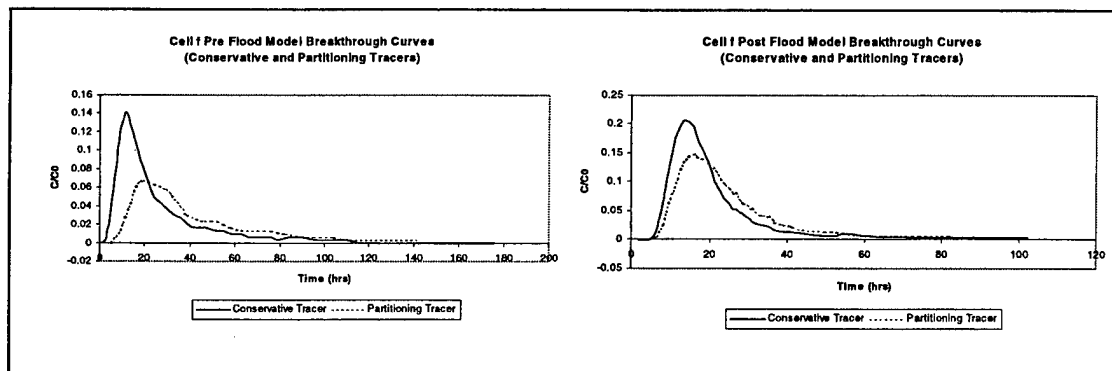Figure 25: Stochastic Model Pre- and Post-Flood Breakthrough Curves (Cell F)

Using moments calculated from these model simulations, the Cell F stochastic

model estimated a cleanup efficiency of approximately 75 % (Table 10). The results

obtained in this model are compared with the moment analysis results in Section 4.4.

Cell ID:	f
EW	3
Model Type	Stochastic
Opt Method	GA

Partitioning Tracer	2,2-dimethyl-3-pentanol
K_{nw}	7.42

Test Type	Layer ID	S_n	Average S_n
Pre Flood	1	0.1115	0.1605
	2	0.1019	
	3	0.2681	
Post Flood	1	0.0544	0.0408
	2	0.0069	
	3	0.0612	
	Cleanup Efficiency		74.57%

Table 10: Stochastic Model Estimate of NAPL Saturation

and Cleanup Efficiency (Cell F)

4.4 Method of Moments and Model Comparisons

The results obtained from the moment analysis of the extraction well tracer data

(extrapolated) are shown in Table 11. The estimates for NAPL saturation and cleanup

efficiency were calculated from the extraction well data extrapolated to account for

concentrations in the tail. Extrapolation was done by performing linear regression

analysis on the plot of the natural log of concentration vs time of the data in the negative

slope portion of the experimental breakthrough curve. The equation for the line (returned

by Excel) was used to estimate concentration for times greater than the time of the final

reported concentration.

Test Cell	Test Type	Tracer Type	Mean Travel Time (hrs)	Est. NAPL Saturation	Est. Cleanup Efficiency
8	Pre Flood	Conservativ	30.14	0.0889	62.10%
		Partitioning	51.94		
	Post Flood	Conservativ	28.97	0.0337	
		Partitioning	36.46		
f	Pre Flood	Conservativ	24.86	0.0861	62.79%
		Partitioning	42.24		
	Post Flood	Conservativ	19.95	0.0320	
		Partitioning	24.84		

Table 11: Method of Moments Results

The comparison of the moment analysis results and the model results for Cell 8 is shown in Table 12, and the comparison for Cell F is shown in Table 13. Differences in estimated saturation levels and cleanup efficiency from inverse modeling and moment analysis can have several sources, including the simplifying assumptions made in the model and fitting errors between the model and experimental data.

Table 12 shows that, despite adjusting the post flood calculations to reflect two layers instead of three, the Cell 8 analytic model overestimated the cleanup efficiency compared to the stochastic model and moment analysis. By observing the scale of the pre-flood and post-flood saturation for all three layers, it is evident that the problem lies in the post-flood analytic model. Analysis of the breakthrough curves for the conservative tracer showed that the model fit the data poorly, as indicated by an increasing Chi Square function in the tail. This error in the tail caused the model to underestimate R and NAPL saturation and explains the high estimate for cleanup efficiency.

Model Type	Test Type	Est. NAPL Saturation	Est. Cleanup Efficiency
Analytic	Pre Flood	0.0785	83.57%
	Post Flood	0.0129	
Stochastic	Pre Flood	0.0997	67.79%
	Post Flood	0.0321	
Moment Analysis	Pre Flood	0.0889	62.10%
	Post Flood	0.0337	

Table 12: Cell 8 Comparison of Results

Table 13 shows the comparison of the Cell F analyses. The stochastic and

analytic models provided similar estimates for pre- and post-flood NAPL saturation, but

both provided significantly higher estimates for the pre-flood saturation than the moment

analysis of the experimental data. This may result from the problems with using a

constant C_0 discussed in Section 3.2.3. Failing to account for possible tracer losses could

result in an overestimation of NAPL saturation. Other simplifying assumptions used in

the models might also contribute to the difference.

Model Type	Test Type	Est. NAPL Saturation	Est. Cleanup Efficiency
Analytic	Pre Flood	0.1558	77.19%
	Post Flood	0.0355	
Stochastic	Pre Flood	0.1605	74.57%
	Post Flood	0.0408	
Moment Analysis	Pre Flood	0.0861	62.79%
	Post Flood	0.0320	

Table 13: Cell F Comparison of Results

Chapter 5. Conclusions and Recommendations for Future Study

5.1 Introduction

This research demonstrated the ability of GAs to perform parameter optimization to model PITT results and estimate NAPL mass and cleanup efficiency. Inverse modeling of PITT results poses a special problem since there are currently no non-linear optimization methods that guarantee convergence to a global optimum for the objective function (Reeves, 1993). This problem is usually addressed through characterization of the aquifer so that the modeler has a good idea of what the optimization parameters may be. This knowledge, combined with multiple model runs using different initial parameter values, can be used to converge on a locally optimum solution that approximates the global optimum. However, if adequate data concerning the hydraulic characteristics of an aquifer are not available, the problems associated with non-linear optimization are compounded. One of the inherent advantages of genetic algorithms are thei ability to "accidentally" discover a more optimum solution. Consequently, iterative application of the GA to inverse modeling of groundwater transport is more likely to converge at a global optimum even if little is known about likely initial values for the decision variables.

5.2 Conclusions

5.2.1 Objective Function

Minimizing the sum of the Chi Square function resulted in a reasonable fit of the model to the experimental data. Attempts to incorporate the number of layers (n) into the model as a decision variable failed due to requirements for conflicting constraints (Section 4.2.1). However, an empirical approach was used to determine a value for n that

met the dual objectives of minimizing model complexity while still obtaining a good fit of the model to the experimental data.

5.2.2 Model Parameters

Use of the conservative tracer data to estimate groundwater flow parameters, and the subsequent use of those parameters to estimate retardation from the partitioning tracer data provided a reasonable fit of the models to the PITT data. The results obtained in this research demonstrated an inverse relationship between the transport velocity of the conservative tracer (as represented by "v" in the analytic model and "mtt" in the stochastic model) and the estimated NAPL saturation for each layer.

5.2.3 Model Results

The NAPL saturation and cleanup efficiencies estimated by the analytic and stochastic models were comparable to results reported by Jawitz *et al.* (1998) and Rao *et al.* (1997) as well as to estimates obtained through moment analysis of the PITT data. The advantages of GAs, as applied in this research, lie in their ability to find a solution that approximates the global optimum for the objective function when little is known about probable values for the decision variables. Iterative application of the GA converges on a good model solution even when the initial parameter values are not close to the values that provide a good model fit.. However, knowledge about probable parameter values greatly reduces the time required for the GA to fit the model to the PITT data.

The results from the post-flood analytic model demonstrated sensitivity to fitting the tail of the breakthrough curve. This problem may be alleviated by fitting the model to extrapolated PITT data, or by increasing the number of model layers.

65

5.3 Recommendations for Future Research

- Heterogeneity in an aquifer, and the corresponding variation in tracer transport times, can result in PITT breakthrough curves that have several peaks and valleys. Inverse modeling can address this, as was done in this research, by optimizing discrete parameter values for multiple layers. Because the extraction well breakthrough curves used in this research fit a simple or bimodal curve, determining the number of layers to obtain a reasonable fit of the model to the experimental data was relatively simple using an empirical approach. However, the PITT data used in this research were obtained from tests conducted in hydraulically isolated test cells, so the complexity of the breakthrough curves was minimized. GAs may provide an appropriate optimization method to balance model complexity (i.e. the number of layers) with the goodness of fit to the PITT data, but this research demonstrated the limitations of Evolutionary Solver at optimizing the number of layers. Including the number of layers as a decision variable for model optimization would require the use of a GA designed to address conflicting constraint requirements.

- The ability of a model to provide a reasonable estimate of NAPL distribution would be enhanced by varying the thickness of the layers. Layer thickness could be incorporated as a decision variable into future models, and the ability of GAs to estimate NAPL saturation could be evaluated.

- Anason (1999) developed a software package in Visual Basic to perform cost-benefit analysis for cosolvent flushing. This cost-benefit tool could be combined with groundwater transport models developed in Excel to provide a comprehensive

66

decision-making tool for site characterization and remediation design, and may provide an avenue for future research efforts.

- The stochastic model developed by Enfield (2000) represents a new way to represent the breakthrough curves obtained from PITTs. This research demonstrated the ability of the stochastic model to represent PITT results obtained under relatively controlled conditions. The applicability of the stochastic model to PITTs conducted on a larger scale should be evaluated in future research.

- Data available from PITTs conducted at Hill AFB include information that was not used in developing the models for this research. This data (e.g., from multilevel samplers, core samples, etc.) could be used, along with GAs, to determine if "real" layers exist that could be modeled or, in other words, to obtain a more reasonable representation of the aquifer heterogeneity and NAPL distribution.

Bibliography

Anason, S. L. Cost-Benefit Analysis of Cosolvent Flushing to Treat Groundwater Contamination Source Areas. MS Thesis, AFIT/GEE/ENV/99M-02, School of Engineering and Environmental Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, Mar 1999.

Armstrong Laboratory (USAF). Groundwater Remediation Field Laboratory. DTIC Report AD-A363 156. Strategic Environmental Research and Development Program, Jun 1997.

Broholm, K., S. Feenstra and J. A. Cherry, "Solvent Release into a Sandy Aquifer. 1. Overview of Source Distribution and Dissolution Behavior." Envir Sci & Tech. Vol 33, No. 5, 681 – 690 (1999).

Brusseau, M. L. and P. S. C. Rao. "Sorption Non-Ideality During Organic Contaminant Transport in Porous Media." CRC Crit. Rev. Envir. Control. Vol. 19, No. 1, 33 – 99 (1989).

Domenico, P. A. and F. W. Schwartz, Physical and Chemical Hydrogeology. New York: John Wiley & Sons, 1998.

Enfield, C., "Tracers Forecast the Performance of NAPL Remediation Projects." submitted for publication to: Groundwater Research. (2000).

Falta, R. W., C. M. Lee, S. E. Brame, E. Roeder, J. T. Coates, C. Wright, A. L. Wood, and C. G. Enfield. "Field Test of High Molecular Weight Alcohol Flushing for Subsurface Nonaqueous Phase Liquid Remediation." Water Resources Research, Vol. 35, No. 7, 2095 – 2108 (Jul, 1999).

Garrett, C. A., J. Huang, M. N. Goltz, and G. B. Lamont. "Parallel Real-Valued Genetic Algorithms for Bioremediation Optimization of TCE-Contaminated Groundwater," Proceedings of 1999 Congress on Evolutionary Computation, Vol 3, 2183 – 2189, Washington, DC, 6-9 July 1999.

Harrouni, K. E., D. Ouazar, G. A. Walters, A. H. D. Cheng. "Groundwater Optimization and Parameter Estimation by Genetic Algorithm and Dual Reciprocity Boundary Element Method." Eng Anal w/ Bound Elem. Vol. 18, 287-296 (1996).

Jawitz, J. W., M. D. Annable, P. S. C. Rao, and R. D. Rhue. "Field Implementation of a Winsor Type I Surfactant/Alcohol Mixture for in Situ Solubilization of a Complex LNAPL as a Single-Phase Microemulsion." Envir Sci & Tech. Vol 32, No. 4, 523-530 (1998).

Jin, M., G. W. Butler, R. E. Jackson, P. E. Mariner, J. F. Pickens, G. A. Pope, C. L. Brown, and D. C. McKinney. "Sensitivity Models and Design Protocol for Partitioning Tracer Tests in Alluvial Aquifers." Ground Water. Vol. 35, No. 6, 964 – 972 (Nov – Dec 1997).

Jin, M., M. Delshad, V. Dwarakanath, D. C. McKinney, G. A. Pope, K. Sepehrnoori, and C. E. Tilburg. "Partitioning Tracer Test for Detection, Estimation, and Remediation Performance Assessment of Subsurface Nonaqueous Phase Liquids." Water Resources Research. Vol. 31, No. 5, 1201 – 1211 (May, 1995).

Lowe, D. F., C. L. Oubre, and C. H. Ward. Surfactants and Cosolvents for NAPL Remediation: A Technology Practices Manual. CRC Press LLC, Boca Raton, FL, 1999.

Lybanon, M. and K. C. Messa, Jr. "Genetic Algorithm Model Fitting." in Practical Handbook of Genetic Algorithms, Complex Coding Systems, Volume III. Ed. Lance Chambers. New York: CRC Press, 1999.

Mercer, M. "Failures and Utility of Groundwater Extraction." n. pag. http://www.epa.gov/epaoswer/hazwaste/ca/naplweb1/Aug1991_files/August1991. htm

Nelson, N. T., M. Oostrom, T. W. Wietsma, and M. L. Brusseau. "Partitioning Tracer Method for the In Situ Measurement of DNAPL Saturation: Influence of Heterogeneity and Sampling Method." Environmental Science and Technology. Vol. 33, No. 22, 4046 – 4053 (1999).

Nelson, N. T and M. L. Brusseau. "Field Study of the Partitioning Tracer Method for Detection of Dense Nonaqueous Phase Liquid in a Trichoroethane-contaminated Aquifer." Environmental Science and Technology. Vol 30, No. 9, 2859 – 2863 (1996).

Rao, P. S. C., M. D. Annable, R. K. Sillan, D. Dai, K. Hatfield, W. D. Graham, A. L. Wood, and C. G. Enfield. "Field-Scale Evaluation of In Situ Cosolvent Flushing for Enhanced Aquifer Remediation." Water Resources Research, Vol. 33, No. 12, 2673-2686 (Dec, 1997).

Reeves, C. R. Modern Heuristic Techniques for Combinatorial Problems. Blackwell Scientific Publications, Osney Mead, Oxford, 1993.

Ritzel, B. J., J. W. Eheart, and S. Ranjithan. "Using Genetic Algorithms to Solve a Multiple Objective Groundwater Pollution Containment Problem." Water Resources Research. Vol. 30, No. 5, 1589 – 1603 (May, 1994).

Sellers, K. Fundamentals of Hazardous Waste Site Remediation. Boca Raton: CRC Press LLC, 1999

Sheely, C. Q. and D. E. Baldwin, Jr. "Single-Well Tracer Tests for Evaluating Chemical Enhanced Oil Recovery Processes." Journal of Petroleum Technology. 1887 – 1895 (Aug, 1982).

Tang, J. S. "Partitioning Tracers and In-Situ Fluid-Saturation Measurements." SPE Formation Evaluation. 33 – 40 (March, 1995).

Trowbridge, B. E., G. Beyke, W. O. Heath, and E. Garcia. "Remediating DNAPLs: New Technology for Treating Dense Non-Aqueous Phase Liquids at DOD and DOE Sites." Environmental Protection. (Jul, 1999).

Van Genuchten, M. Th. and W. J. Alves, Analytical Solutions of the One-Dimensional Convective-Dispersive Solute Transport Equation. US Dept of Ag, Ag Research Sevice, Tech Bulletin # 1661, 1982.

Wang, M. and C. Zheng. "Optimal Remediation Policy Selection under General Conditions." Groundwater. Vol. 35, No. 5, 757 – 764 (Sep – Oct, 1997).

Young, C. M., R. E. Jackson, M. Jin, J. T. Londergan, P. E. Mariner, G. A. Pope, F. J. Anderson, and T. Houk. "Characterization of a TCE DNAPL Zone in Aluvium by Partitioning Tracers." GWMR. (Winter, 1999).