Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Contaminant Hydrology 67 (2003) 247–268
JOURNAL OF Contaminant Hydrology
www.elsevier.com/locate/jconhyd

The effect of entrapped nonaqueous phase liquids on tracer transport in heterogeneous porous media: laboratory experiments at the intermediate scale

Gilbert R. Barth[a,b,*], Tissa H. Illangasekare[c], Harihar Rajaram[a]

[a] *Department of Civil and Environmental Engineering, University of Colorado, Boulder, CO 80309-0428, USA*
[b] *U.S. Geological Survey, 3215 Marine Street, Boulder, CO, USA*
[c] *Division of Environmental Sciences and Engineering, Colorado School of Mines, Golden, CO 80401-1887, USA*

Received 15 February 2002; accepted 19 February 2003

Abstract

This work considers the applicability of conservative tracers for detecting high-saturation nonaqueous-phase liquid (NAPL) entrapment in heterogeneous systems. For this purpose, a series of experiments and simulations was performed using a two-dimensional heterogeneous system (10 × 1.2 m), which represents an intermediate scale between laboratory and field scales. Tracer tests performed prior to injecting the NAPL provide the baseline response of the heterogeneous porous medium. Two NAPL spill experiments were performed and the entrapped-NAPL saturation distribution measured in detail using a gamma-ray attenuation system. Tracer tests following each of the NAPL spills produced breakthrough curves (BTCs) reflecting the impact of entrapped NAPL on conservative transport. To evaluate significance, the impact of NAPL entrapment on the conservative-tracer breakthrough curves was compared to simulated breakthrough curve variability for different realizations of the heterogeneous distribution. Analysis of the results reveals that the NAPL entrapment has a significant impact on the temporal moments of conservative-tracer breakthrough curves.

Keywords: Nonaqueous phase liquids; High-saturation NAPL entrapment; Heterogeneous systems; Tracer tests; Gamma-ray attenuation system; Temporal moments

* Corresponding author. Currently at: Waterstone, 1650 38th Street, Suite 201E, Boulder, CO 80301, USA. Tel.: +1-303-444-1000.
E-mail addresses: barth@waterstoneinc.com (G.R. Barth), tillanga@Mines.EDU (T.H. Illangasekare), hari@bechtel.colorado.edu (H. Rajaram).

0169-7722/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0169-7722(03)00066-4

1. Introduction

Chlorinated solvents and petroleum products in their organic liquid forms are pervasive in industrial applications. As a result of improper disposal and accidental releases, these nonaqueous-phase liquids (NAPLs) comprise a significant source of ground-water contamination. Laboratory, field studies and numerical simulations (e.g., Illangasekare et al., 1995a,b; Kueper and Frind, 1991; Kueper and Gerhard, 1995) suggest that NAPL distribution in heterogeneous aquifers is highly complex and leads to relatively immobile NAPL at saturations much higher than residual in coarser lenses. This phenomenon is often referred to as macroscale entrapment. The presence of entrapped NAPL at relatively high local saturations changes local permeability (e.g., Illangasekare et al., 1995b; Brooks and Corey, 1966). This paper focuses on assessing the impact of macroscale entrapped NAPL on conservative tracers using laboratory experiments and numerical simulations.

Characterization of NAPL distribution in the subsurface is required for efficient remediation. Characterization techniques often consist of core sample evaluation, cone penetrometer testing, or geophysical logging. Values determined with these methods are subject to large uncertainties because of the highly variable nature of NAPL distributions in heterogeneous media. Mayer and Miller (1992) presented evidence that soil sample volumes required for assessing NAPL saturation are much larger than volumes analyzed for core samples and geophysical logs.

As an alternative, tracer tests across the suspected contamination zone have been proposed for detection of entrapped NAPL. Tracer test results reflect an integrated response of the subsurface overcoming the inherent limitation of discrete core-scale sampling in heterogeneous systems. Tracer tests have become a common method of determining parameter values such as hydraulic conductivity (K) and dispersivity (Adams and Gelhar, 1992; Freyberg, 1986; Garabedian et al., 1991; Moltyaner and Wills, 1991; Sudicky, 1986). Well-documented field tracer tests have been conducted at the Borden Air Force Base in Canada (Freyberg, 1986; Mackay et al., 1986) and the Otis Air Force Base at Cape Cod in the United States (Garabedian et al., 1991; LeBlanc et al., 1991). These tests used a variety of nonreactive (e.g., Chloride, Bromide) and reactive (e.g., Lithium, Fluoride, Carbon Tetrachloride) tracers and a large number of samplers to capture the spatial distribution of the tracer plume as it changed with time. Analysis of the concentration data produced values for hydraulic conductivity and dispersivity at scales relevant to the tests (Freyberg, 1986; Garabedian et al., 1991).

Spatial characterization of tracer plumes requires an extensive network of samplers that is impractical for most situations. Most of the works cited above, which were based on spatial analysis of tracer plume movement, are exceptions because of their extensive network of samplers in three dimensions. A more common field-site investigation would be limited to a sampling-well transect, perpendicular to the overall flow direction at the site, collecting temporal concentration data as the tracer plume passes the plane of sampling wells. Although less common, temporal moment interpretation of tracer data has been performed on observations from several field sites. The work of Dagan et al. (1997) and Rubin and Ezzedine (1997) demonstrated the capability of basic hydrogeological characterization using temporal moments of breakthrough curves (BTCs).

Jin et al. (1995) introduced partitioning interwell tracer tests (PITT) as a method for detecting and estimating residual NAPL in the subsurface using multiple tracers. The method evaluates differences in the behavior of conservative and partitioning tracers as a function of the presence of residual NAPL. In order to develop the relationship between partitioning tracer retardation and the quantity of NAPL present, the approach assumes a homogeneous residual saturation for the region evaluated by the partitioning tracer. Partitioning coefficients are estimated using one-dimensional laboratory column evaluations of a tracer's breakthrough compared to a conservative tracer. The difference between conservative and partitioning transport is reflected by the first temporal moment of the tracer BTCs. The assumption is that with increasing NAPL interfacial area a partitioning tracer is increasingly retarded, resulting in a larger first moment. First moments estimated from BTCs of conservative tracers are intended to reflect transport unaffected by the presence of NAPL and are used for comparison to the first moment of the partitioning tracer BTC. James et al. (2000) used nonlinear regression and simulated partitioning-tracer BTCs to determine the distribution of residual NAPL saturations and the relative efficiency of several sampler configurations. Jin et al. (1995) and James et al. (2000) demonstrate the potential for assessing residual NAPL distributions using tracer BTCs. Their work, however, does not address higher-than-residual NAPL saturations, i.e., macroscale entrapment, and its potential for impacting conservative tracer transport.

This paper addresses the question: does macroscale entrapped NAPL have a significant impact on conservative transport in a heterogeneous, two-dimensional system? The approach envisaged consists of comparing conservative tracer test results from a pristine region of an aquifer to results from a region of the same aquifer where an NAPL spill is suspected. Will tracer test results from the spill zone be significantly different from the pristine portion of the aquifer? The approach requires an aquifer with statistically stationary heterogeneity over a large enough region. Observations at the Borden and Cape Cod sites suggest that this may not be an unreasonable assumption given that most NAPL contamination regions have lateral dimensions on the order of only a few tens of meters. In order to work with an explicitly characterized heterogeneity, this paper uses intermediate-scale laboratory experiments and numerical simulations to examine the impact of macroscale entrapment on conservative-tracer tests. The term intermediate-scale refers to the overall size of the experiment, distinguishes the laboratory experiments presented from typical laboratory experiments, and should not be confused with the term macroscale entrapment. As discussed above, macroscale entrapment denotes immobile NAPL at high saturations. Laboratory experiments are typically limited in terms of size and complexity, often consisting of one-dimensional, homogeneous systems, or systems with very simple heterogeneity. In contrast, the experiments presented in this paper were performed in a 10-m-long tank with a relatively complex heterogeneous porous medium. The combination of physical and numerical experiments provides insight to the capabilities of tracer techniques for the characterization of subsurface materials.

To achieve the objectives of this paper, a heterogeneous system was created and instrumented. Tracer tests were designed, conducted and analyzed. A series of NAPL spills were performed. Qualitative and quantitative monitoring of the NAPL spill

experiments documented the temporal and spatial distribution and saturation of NAPL within individual lenses in the heterogeneous system and provided insight to the behavior of NAPLs in heterogeneous systems. Analysis of the tracer-test BTCs quantified the impact of macroscale entrapped NAPL on a conservative tracer. Simulated tracer experiments using different realizations of the heterogeneous system provided insight to transport variability between realizations and were used to evaluate the significance of the macroscale entrapped NAPL on conservative-tracer transport.

2. Intermediate-scale experimental system

2.1. Construction and instrumentation of the heterogeneous porous medium

The intermediate-scale porous medium was constructed in a tank approximately 10 m long, 1.2 m tall, and 0.06 m inside width (Fig. 1a). Each end of the porous medium consisted of a 20-cm section of pea gravel to provide constant-head boundaries for the system. The overall gradient and saturated zone thickness were adjusted with two constant-head tanks that controlled the water level in the pea gravel. The water table was level with the top of the sand packing at the up-gradient end of the tank. At the down-gradient end it was 15.7 cm below the top of the packing, producing an overall gradient of approximately 0.016 along the length of the tank. Deionized water was supplied to the up-gradient constant-head tank. The gradient and resulting flow of approximately 3.2 l/h were maintained throughout the experiments.

The packing within the tank consisted of two zones: a homogeneous section of coarse sand (#8 sieve) in the upstream 1.1 m of the tank, followed by an 8.1-m heterogeneous section (Fig. 1a). The heterogeneous zone served as a simplified laboratory analogue of random field-site sedimentary structure, was created using five different sands, and designed to support explicit representation in a numerical model. The heterogeneous zone approximated a log-normal distribution of hydraulic conductivity (K) with a mean ln$K$ value of 4.18 ($\mu_{\ln K}$), where K has units of cm/h, a variance ($\sigma^2_{\ln K}$) of 1.22 and an anisotropic exponential covariance structure. For comparison to well-known field sites, $\sigma^2_{\ln K}$ of the Cape Cod, Borden and Columbus sites were 0.2 (Hess et al., 1992), 0.24 (Woodbury and Sudicky, 1991) and 4.5 (Boggs et al., 1990), respectively. The horizontal (λ_h) and vertical (λ_v) correlation scales of ln$K$ were 50.8 and 5.08 cm, respectively, leading to an anisotropy ratio, $\lambda_v/\lambda_h = 0.1$. At the Cape Cod, Borden and Columbus sites, $\lambda_v/\lambda_h = 0.04$, 0.02 and 0.01, respectively. Tank size required limiting λ_v/λ_h to about 2–10 times these field values but still allowed creation of a system with significant anisotropy, heterogeneity, and the resulting potential for macroscale entrapment.

To design the heterogeneous packing, a continuous ln$K$ distribution with the desired statistical structure was generated using a Fourier summation algorithm (Shinozuka and Jan, 1972) and then discretized into five categories. Each category was assigned a particular sieve size sand: #16, #30, #50, #70 or #110 (Table 1). Permeameter measured hydraulic-conductivity values (K_p in Table 1) for each sieve size sand were determined by Mapa et al. (1994) (unpublished report for the U.S. Army Engineer Waterways

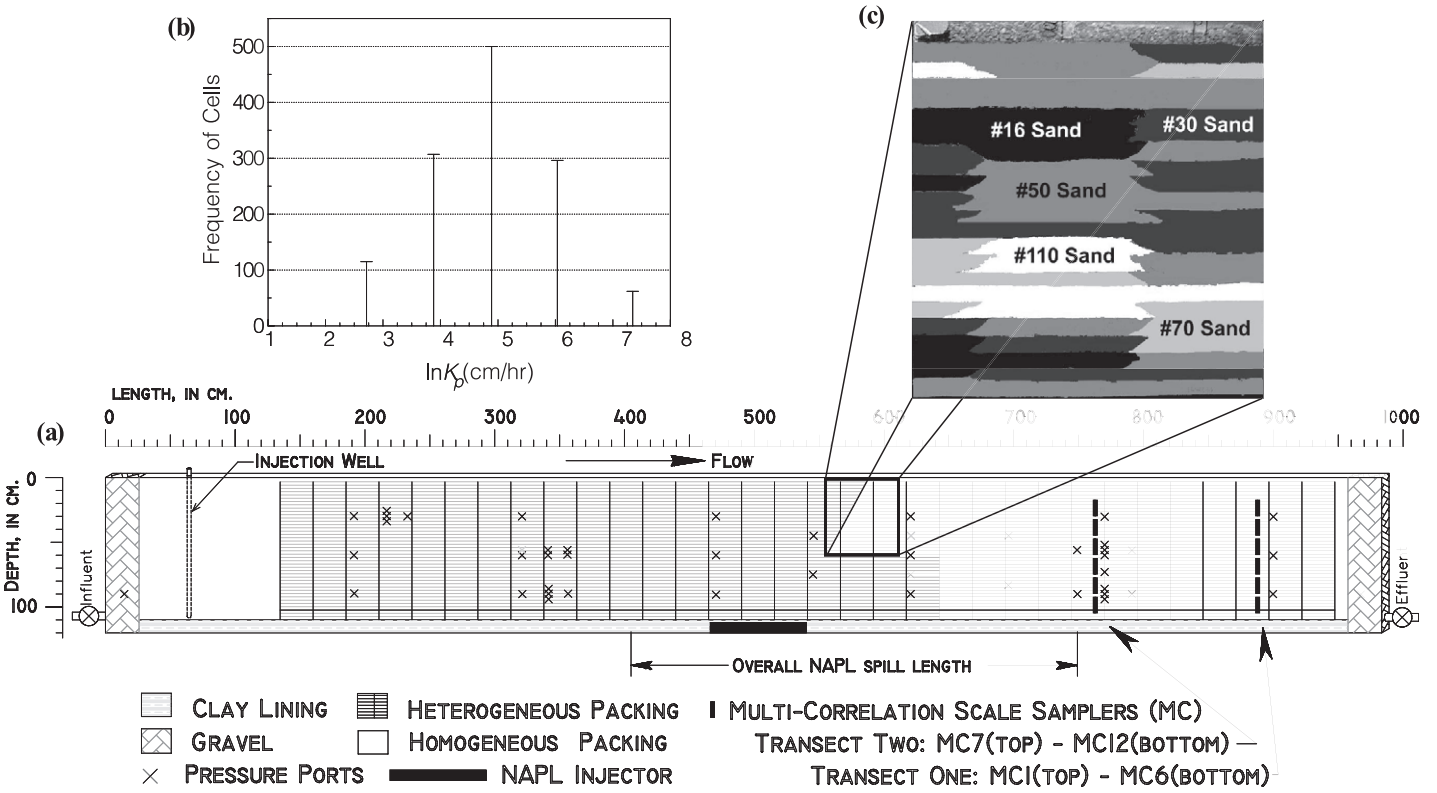


Fig. 1. Intermediate-scale tank design showing (a) the dimensions of the entire tank including a simplified representation of the sand cell packing and the pressure and sampling ports, (b) the frequency distribution of sand cell packing for the five hydraulic conductivities, and (c) an enhanced gray-scale detail of the tank packing. Each sand lens is 25.4-cm long and 2.54-cm tall, and is tapered at the ends as shown in (c).

Table 1
Sand properties

	Mesh size (ASTM E-11)[a]				
	#16	#30	#50	#70	#110
K_p (cm/h)	1550	417	133	48.6	15.1
K_r (cm/h)	3170	716	156	104	45.1
d_{50} (mm)	0.88	0.49	0.30	0.19	0.103
d_{60}/d_{10}	1.72	1.50	1.94	1.86	NA
λ[b]	3.5	3.0	1.9	2.2	NA
S_r	0.07	0.26	0.28	0.30	0.30

[a] From American Society for Testing and Materials (1994).
[b] From Szlag (1997).

Experiment Station). The values of $\mu_{\ln K}$ and $\sigma^2_{\ln K}$ for the discretized distribution are as noted above. The permeameter-measured frequency distribution of the five different sands in the heterogeneous packing is shown in Fig. 1b. The homogeneous zone provided a region to inject the tracer and promote initial mixing as it exited the injection well, producing a relatively consistent vertical line source. The coarser sand, with a relatively high dispersivity, reduced the effect of micro-heterogeneities in the packing and the potential for variation from the injection well.

A consistent packing procedure was used for the entire tank. The sand was wet-packed in the tank to minimize consolidation and air entrapment. Each cell measured 25.4-cm long and 2.54-cm tall with overlapping ends (Fig. 1c). Vertical interfaces were avoided to reduce the chance for preferential migration of NAPLs. A total of 1280 cells were packed: 32 columns and 40 layers producing 16 lateral and 20 vertical correlation scales. Packing of the tank followed a set procedure for each layer. First, using the constant-head tanks, the static water table was raised to approximately 4 cm above the most recently packed layer, or the base of the tank for the first layer. Depending on the layer's packing, up to 33 paddles, one on each edge of a sand cell, were mounted along the length of the tank, hanging down into the tank, providing separation between adjacent cells of different sands. Using a consistent length of drop tube and filling in the order from finest to coarsest, sands were dropped into their respective cells. After an individual or group of cells consisting of the same sand were filled, and prior to filling the neighboring coarser sands, the paddles were removed to allow the ends of the sand cell to slump so that the interfaces between dissimilar sands were not vertical (Fig. 1c).

Tracer was injected using a specially designed injection well. The well consisted of a 0.635-cm diameter stainless-steel tubing injector with 24 evenly spaced outlets covered by #200 stainless-steel mesh inserted into a fully screened, 2.54-cm diameter PVC casing. Packers mounted on the stainless steel tubing ensured delivery of the tracer into the casing over the injection interval. The line source injected for the reported experiments was 36-cm tall with the top located at a depth of 34 cm. The vertical source interval ensured that, despite relatively small vertical dispersion, the plume's vertical dimension spanned multiple correlation scales when it entered the heterogeneous section. The resulting tracer plume reflected the overall transport characteristics

of the hydraulic-conductivity distribution rather than a specific portion of the packed realization. Dye injections were performed in the tank to verify creation of a line source that was uniform both in the vertical direction and across the width of the tank. However, creation of a tracer plume with sufficient visual contrast to document its progress throughout the 10-m long tank was outside the objectives of this investigation.

Samplers were designed to reduce the potential for local hydraulic conductivity to dominate sample constitution (Barth, 1999). Two transects, each with six samplers, were used to measure BTCs (Fig. 1a). The samplers are designed to provide a sample across approximately two vertical correlation scales. They require far less purge volume than a well installed from the top of the packing, thereby minimizing the potential impact of sampling on the flow field. This is especially important given the potential for disrupting the flow field. Each sampler consists of a vertical 10-cm long perforated 0.32-cm diameter piece of copper tubing wrapped with #200 stainless steel mesh. One end of the tubing is sealed and the other bent 90° and attached to a bulkhead compression fitting, allowing sampling through the tank wall. The sampler requires extraction of only about 1.5 ml to provide a 1-ml sample. Two transects of six samplers each (Fig. 1a), labeled MC1 (top)–MC6 (bottom) and MC 7 (top)–MC12 (bottom) for transects 1 and 2, respectively, were installed. The samplers cover all but the top 10 cm and bottom 5 cm of the porous media. The saturated thickness above MC1 and MC7 is on the order of a few centimeters due to the location of the sampler transect along the tank (Fig. 1a) in relation to the sloping water table. The sampler interval, due to its length, did not correspond to an exact integer number of sand cells. In addition, neither the top nor the bottom of each sampler was positioned exactly at the interfaces between sand cells.

2.2. Tracer experiments prior to spilling NAPL

A total of six tracer tests was performed using either bromide (KBr) or tritium as the conservative tracer (Table 2). Tests C7, C8 and C9 used bromide, and D1, D2 and D4 used tritium tests. The switch from bromide to tritium was made when it became apparent that entrapped NAPL might result in significantly lower local gradients and the potential for density-induced sinking (Barth et al., 2001a). For each experiment the

Table 2
Tracer injections

Experiment name	Flow rate through tank (l/h)	Injection			*C* (molarity or dpm)
		Interval (h)	Volume (l)	Tracer	
C7	3.31	0.91	2.8	bromide	0.002
C8	3.26	0.92	2.7	bromide	0.002
C9	3.03	0.92	2.8	bromide	0.002
D1	3.30	0.93	2.8	tritium	2.94e+08
D2	2.68	1.02	2.8	tritium	2.94e+08
D4	2.26	1.32	2.8	tritium	2.94e+08

injection rate was approximately 3.0 l/h, just slightly less than the nominal tank effluent rate, to avoid flow field disruption. Samples were collected every 4 h, or approximately every 0.08 pore volume until roughly 3.5 pore volumes had passed through the tank. Samples from the KBr experiments were analyzed using an ion selective electrode. A liquid scintillation counter was used to analyze the tritium samples. The first four tracer tests, C7, C8, C9 and D1, were very similar. They were all performed prior to spilling any NAPL and, except for an approximately 10% decrease in flow through the tank for C9, all other experimental conditions were maintained within their respective margins of experimental error. Comparison of breakthrough curves from C7, C8 and D1 provides an indication of the level of reproducibility (Fig. 2). Considering the similarity between these results, BTCs from C7 are used to represent pre-NAPL spill transport. The C9 test illustrates that the system responds appropriately for small changes in boundary conditions; in this case, a decrease in the overall gradient.

Breakthrough curves can be used to assess the impact of a particular distribution of K on tracer transport through a system (e.g., Dagan et al., 1997; Rubin and Ezzedine, 1997). Experiments C7, C8 and D1 were used to ascertain the transport characteristics of the heterogeneous distribution of *K* prior to any NAPL entrapment. The tracer tests provided BTCs from each sampler. The BTCs from each sampler were combined using flow-weighted concentrations, producing a single BTC for each transect from each tracer test (Barth, 1999). Flow values for weighting the concentrations were obtained from numerical simulations calibrated using head and flow data (Barth et al., 2001b). Calculation of the BTC temporal moments provided a quantitative summary of

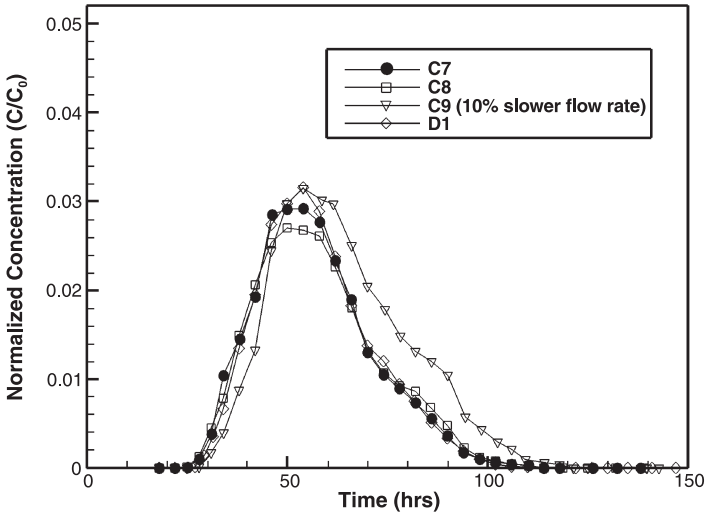


Fig. 2. Conservative tracer breakthrough curves prior to spilling NAPL. Tests C7, C8 and C9 used bromide; D1 used tritium. Flow was about 10% slower for C9. Comparison of C7, C8 and D1 indicates a high level of repeatability. Experiment C9 demonstrates reasonable system response to boundary condition changes.

conservative transport characteristics for the heterogeneity (e.g., Dagan et al., 1997; Rubin and Ezzedine, 1997). The jth absolute time moment (M_j) is defined as:

$$M_j = \int_0^\infty t^j C(x,t)\mathrm{d}t \tag{1}$$

where t is time and $C(x,t)$ is concentration as a function of space and time. The normalized absolute jth moment (m_j) is obtained by dividing M_j by M_0,

$$m_j = \frac{M_j}{M_0} \tag{2}$$

and μ_j represents the jth normalized central moment.

$$\mu_j = \frac{\int_0^\infty (t - m_1)^j C(x,t)\mathrm{d}t}{M_0} \tag{3}$$

In this paper m_1 and μ_2 are evaluated for each BTC, serving as lumped indicators of tracer transport characteristics, and are referred to simply as the first (m_1) and second (μ_2) moments, respectively. The first moment indicates the mean arrival time of the tracer at the sampler and μ_2 quantifies spreading of the plume and provides an indication of effective dispersivity.

3. NAPL-spill experiments

The NAPL spill experiments were designed to allow capillary and buoyancy forces to redistribute an NAPL in a manner similar to DNAPL sinking below the water table. Unfortunately, most of the DNAPLs available for laboratory studies are fairly toxic and would tend to corrode the walls of the Plexi-glass lined intermediate-scale tank. An aliphatic hydrocarbon, Soltrol 220 (Phillips Petroleum), was selected as the primary NAPL component for injection based on its relatively low toxicity. However, Soltrol is an LNAPL; it can be considered an experimental analog of fuel oil and has a density of 0.81 gm cm^{-3} (Table 3). The interfacial tension of Soltrol 220 and water is 23.45 dyn cm^{-1} (Dai et al., 2001). In order to produce a scenario analogous to DNAPL

Table 3
Fluid properties

Fluid	Density (gm/cm^3)	Lumped calibration constants: U (cm^{-1})
Soltrol	0.81	NA
1-Iodoheptane	1.38	NA
Soltrol mixture	0.83	0.716
Water	1.00	0.186

penetrating downward into the saturated zone, Soltrol was injected from the bottom of the water-saturated heterogeneous packing. Density contrasts between the water and Soltrol forced the Soltrol upward in the same manner as a DNAPL would penetrate downward through the saturated zone.

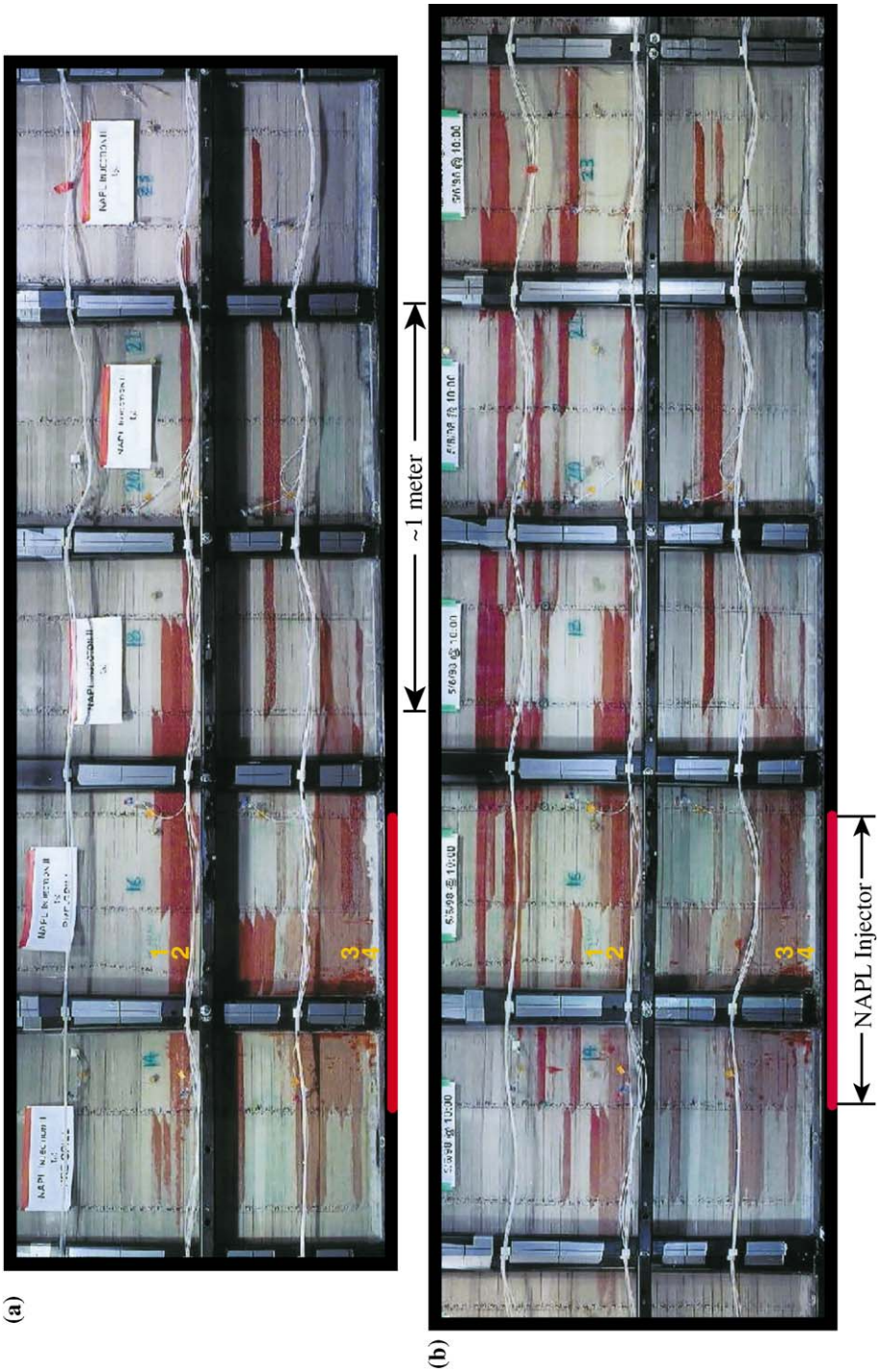
The Soltrol was mixed with 1-Iodoheptane at a 1:9 ratio to produce a NAPL mixture with a gamma-particle attenuation coefficient significantly greater than water and density still significantly less than water ($\rho = 0.83$ gm cm^{-3}). Automate Red B (Morton International), a dark red organic dye essentially insoluble in water, was added to the Soltrol/1-Iodoheptane mixture at a ratio of 0.1% (volume/volume) to allow visualization of the spill. For the remainder of the paper the mixture of Soltrol 220, 1-Iodoheptane and Automate Red B is simply referred to as the NAPL.

The NAPL injector consisted of a perforated PVC pipe wrapped with a stainless steel mesh and packed in #8 sieve size sand below columns 14, 15 and 16 of the heterogeneous packing (Fig. 1a) producing a source of 1.5 λ_h, or 76.2-cm long. The NAPL was injected at a rate of about 0.208 and 0.131 l/h for the first and second spill, respectively. The length of injector and slow injection rate promoted lateral spreading of the NAPL (Kueper and Gerhard, 1995) and preferential infiltration into coarser material (Compos et al., 1998) resulting in a spill distribution similar to that produced by a slow leak from a DNAPL tank (Fig. 3). The first spill was 7.2 l and an additional 4.2 l was injected for the second spill. These spill volumes corresponded to 7% and 4% of the total pore volume in the tank. Conservative tracer tests D2 and D4 were performed following the first and second NAPL spills, respectively, and were initiated only after gamma scans, discussed below, indicated that the rapid redistribution of NAPL immediately following injection had ceased. Concentration observations from D2 and D4 were used to calculate the temporal moments of BTCs impacted by the NAPL spill.

3.1. Mapping the NAPL distribution

This study evaluated a two-component system within the saturated zone: water and NAPL. A gamma system was used to measure NAPL saturations during and after the NAPL spills. The gamma system was used in two ways: (1) monitoring the rate of NAPL redistribution immediately following each spill, and (2) mapping of the NAPL saturation distribution during each of the conservative tracer tests. Resolution of the gamma system was about 2.4 cm^2. A detailed description of the system can be found in Illangasekare et al. (1994) and Walser (1995). The scans provided an independent, nonintrusive, quantitative mapping of NAPL saturations. As discussed below, the NAPL saturation data were used to calculate changes in local wetting-phase (water) permeability as a result of NAPL entrapment.

Fig. 3. Digital images of NAPL entrapped in the intermediate-scale heterogeneous porous media. Images were taken after rapid redistribution had ceased. (a) After first spill of 7.2 l. (b) After second spill of an additional 4.35 l. The scales on the two images are approximately the same. Numbers in yellow indicate the scanning locations used in Fig. 4.


(a)
(b)
~1 meter
NAPL Injector

3.2. Characterization of the NAPL spills

Color digital images of the first and second spill provide a qualitative indication of NAPL distribution for each spill and illustrate the dramatic contrasts in NAPL saturations and the complex pattern of spatial distribution (Fig. 3). Comparisons between the two sides of the tank revealed only minor differences in the visual appearance of the NAPL distribution, occurring primarily when NAPL first penetrated into a sand lens. The images in Fig. 3 correspond to just over one third of the entire tank length. Over a 14-day period, the first NAPL spill spread approximately 50 cm in the vertical direction and 300 cm laterally. While these values summarize the extent of NAPL spreading, they do not capture the highly discontinuous nature of the saturation distribution and the dramatic local contrast in NAPL entrapment that was visibly apparent and confirmed using the gamma scans. The majority of redistribution occurred during the first 4 days (Fig. 4). It should be noted that, following the first spill, the entire upper half of the tank remained free of NAPL. The first spill only reached coarse lenses in the bottom half of the tank. The second spill almost eliminated the possibility for advective transport to bypass the entire NAPL spill region: all coarse lenses, from the top to the bottom of the tank, in the spill-impacted region were saturated with NAPL.

Photographic and gamma scan evidence suggest that NAPL from the second spill flowed through the same cells as in the first spill until most of the cells infiltrated during

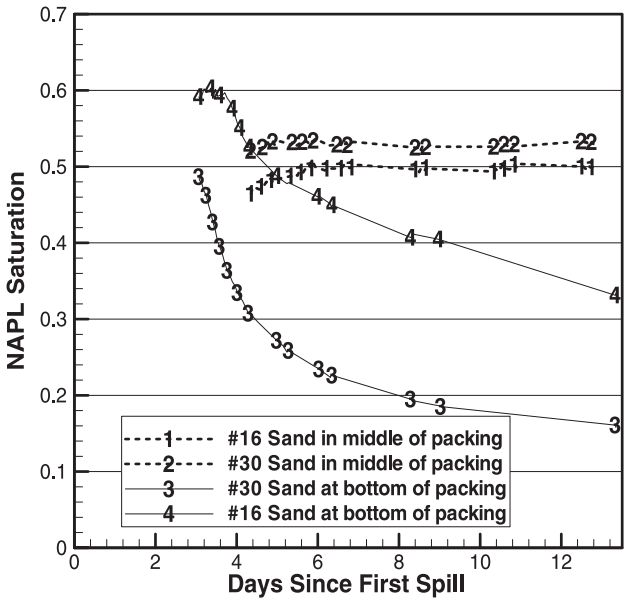


Fig. 4. Change in NAPL saturation with time at four locations, 1, 2, 3, and 4 indicated in Fig. 3, following the first spill. Lower cells drained as buoyancy drove the NAPL higher in the packing. Higher cells, at the upper extent of the spill, exhibited a relatively constant saturation during the period of scanning.

the first spill were fully resaturated. After that, the NAPL began infiltrating new cells at the periphery of the first spill, moving preferentially into coarser material. Conditions during the second spill resulted in infiltration into new cells not occurring until some time after pumping had ceased: only buoyancy and capillarity contributed to infiltration of new cells during the second spill. NAPL migration jumped vertically to coarse lenses and then followed them laterally for large distances. In general, the additional cells infiltrated during the second spill were more widely dispersed than the cells infiltrated during the first spill. The first-spill region acted as a large injector, allowing the spill progress to selectively infiltrate into a sampling of cells more closely resembling the distribution as a whole.

A notable feature of both spills is the apparently disconnected distribution of NAPL-saturated lenses. Medium- to fine-grained lenses were free of virtually any NAPL. The absence of any visual or gamma-scan indication of NAPL in the fine cells separating the NAPL-saturated coarse lenses suggests that pore-scale fingering was the primary means by which the NAPL moved between coarse lenses. Within coarse lenses, as NAPL saturation and the capillary pressure increased, NAPL penetrated the overlying finer lens through only the largest pores. Initiation of NAPL fingers between nonadjacent coarse lenses provided sufficient opportunity for the NAPL to progress between coarse lenses so that the capillary pressure never exceeded the average pore-entry pressure of the finer sands. In this way, NAPL transport and the resulting distribution shown in Fig. 3 were determined by variability at the pore scale.

The impacts of the NAPL spill on the hydraulic behavior of the medium can be quantified based on the modified wetting phase permeability resulting from NAPL entrapment. Although relative permeability was not measured directly, permeability modifications resulting from NAPL entrapment were estimated using the measured NAPL saturations and the Brooks–Corey relative permeability function:

$$k_{rw} = S_e^{(2+3\lambda)/\lambda} \tag{4}$$

where λ is a pore size distribution index, S_e is defined as

$$S_e = \frac{S_w - S_r}{S_m - S_r} \tag{5}$$

and S_w, S_m and S_r are the wetting-phase, maximum and residual saturations, respectively. For laboratory conditions, with relatively uniform media carefully saturated, the value of S_m is essentially equal to 1.0. The values of λ and S_r corresponding to each sand were obtained from suction-saturation measurements by Szlag (1997) and are listed in Table 1. The modified ln$K$ distributions within the NAPL spill zone, reflecting the influence of NAPL entrapment, are shown in Fig. 5. The frequency of large ln$K$ values is reduced in the modified distributions reflecting the tendency of the NAPL to occupy coarse sand regions. The spread of the distribution increases dramatically due to the very low hydraulic conductivity of the lenses with macroscale NAPL entrapment. The variograms of the modified ln$K$ distribution in Fig. 6 illustrate the change in spatial structure of the ln$K$ variation resulting from NAPL entrapment. The variogram sill reflects a dramatic increase in $\sigma^2_{\ln K}$ after the first spill, and only a slight increase

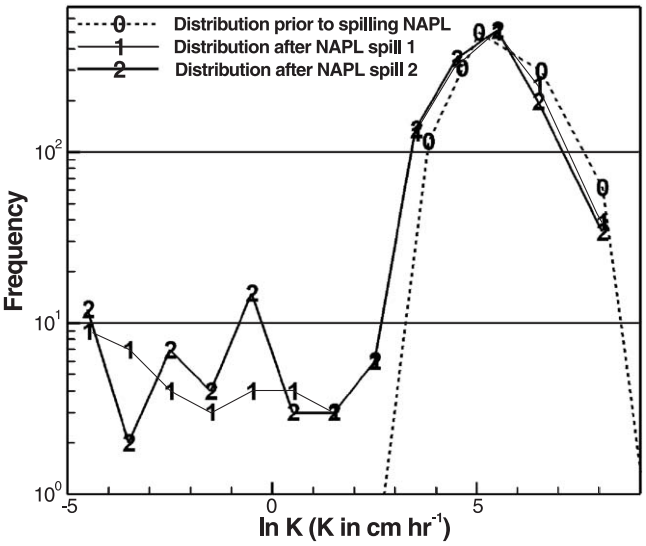


Fig. 5. Initial (K_r) and NAPL modified distribution of ln$K$ following each NAPL spill.

between the first and second spills. The shapes of the variograms are quite interesting. In the case of the horizontal variogram, the major difference between the pre-spill and post-spill variograms occurs at smaller separations. At separations greater than the visually and gamma-scan-determined spill extent, the post-spill variograms approach the

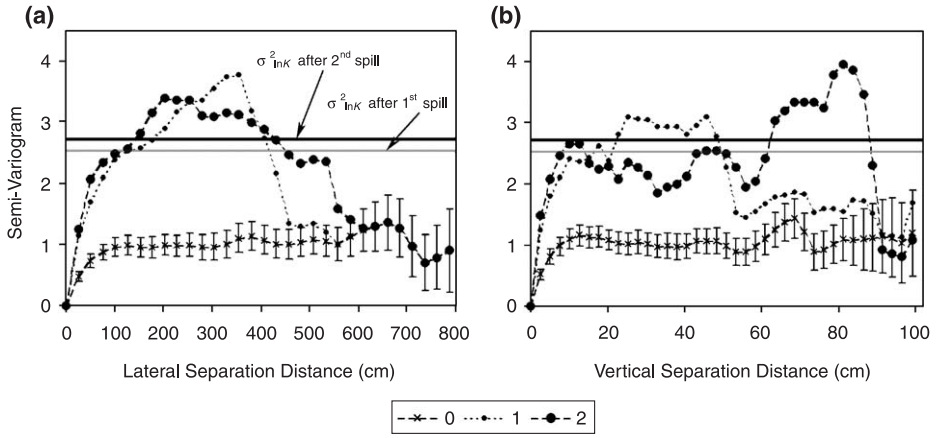


Fig. 6. Variograms of the heterogeneous porous media. Each graph includes variograms based on: the negative exponential covariance function using K_r values prior to any spill (0), with 95% confidence intervals indicated, and the distribution of effective permeability following the first (1), and second NAPL spill (2). (a) Horizontal variograms. Structure from the first and second spill persists for ~ 400 and 550 cm, respectively. (b) Vertical variograms. Structure from the second spill extends almost twice as far as the first spill. Large first spill increase in sill and minimal change for second spill reflect the change in distribution variance with each spill.

general trend of the pre-spill variogram. This reflects the finite horizontal extent of the spill region and the lateral extent of lenses with permeability modified by macroscale entrapped NAPL. The range of separations over which the pre-spill horizontal variogram is modified increases after the second spill, consistent with the increase in the horizontal extent of the spill. The vertical variogram exhibits a similar response for the first spill, but, following the second spill, is very different from the pre-spill vertical variogram over the entire range of separations. This is because the NAPL extends over almost the entire vertical extent of the tank (Fig. 3b).

4. Post-spill flow and transport

Tracer experiments and flow measurements characterizing the NAPL-impacted system commenced once rapid redistribution of the NAPL ceased, typically about 6 days after the NAPL was injected (Fig. 4). Boundary conditions for tracer tests D2 and D4 were the same as C7, C8 and D1. As expected, the NAPL spills affected the bulk flow through the tank. Summary measures such as flow rate and equivalent K, which do not account for the spatial relation of materials, exhibited a consistent response with increasing NAPL entrapment. Compared to pre-spill values the equivalent K, based on effluent rate and overall gradient, was reduced by 19.5% and 32.2% after the first and second spills, respectively. The incremental decrease in K per liter of NAPL for the first spill ($\sim$ 2.9%) was similar to that for the second spill ($\sim$ 3.4%). The fairly consistent incremental decrease is not unexpected: it reflects the combination of material properties, their

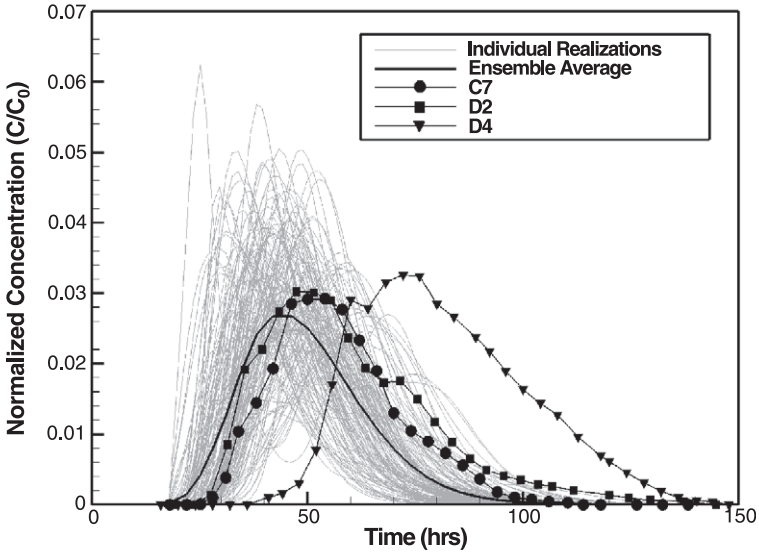


Fig. 7. Conservative-solute BTCs for transect 2. Gray lines are simulated results from 150 realizations of the heterogeneous distribution. The solid line represents the ensemble average of the numerical simulations. Also depicted are observed results prior to the NAPL spill (C7), after the first spill (D2), and after the second spill (D4).

stationary distribution, fluid properties, and the propensity for NAPL to saturate the coarser material.

Qualitatively, modification of the pre-spill BTC by the first NAPL spill appears to be small (Fig. 7). Temporal moments of the observed BTCs (Fig. 8) provide a quantitative assessment of the entrapped NAPL's impact as a function of spill volume. Following the first spill the first moment does not change substantially, while the second spill appears to have a much more dramatic impact. The second moment exhibits much greater initial sensitivity to the presence of NAPL, increasing consid-

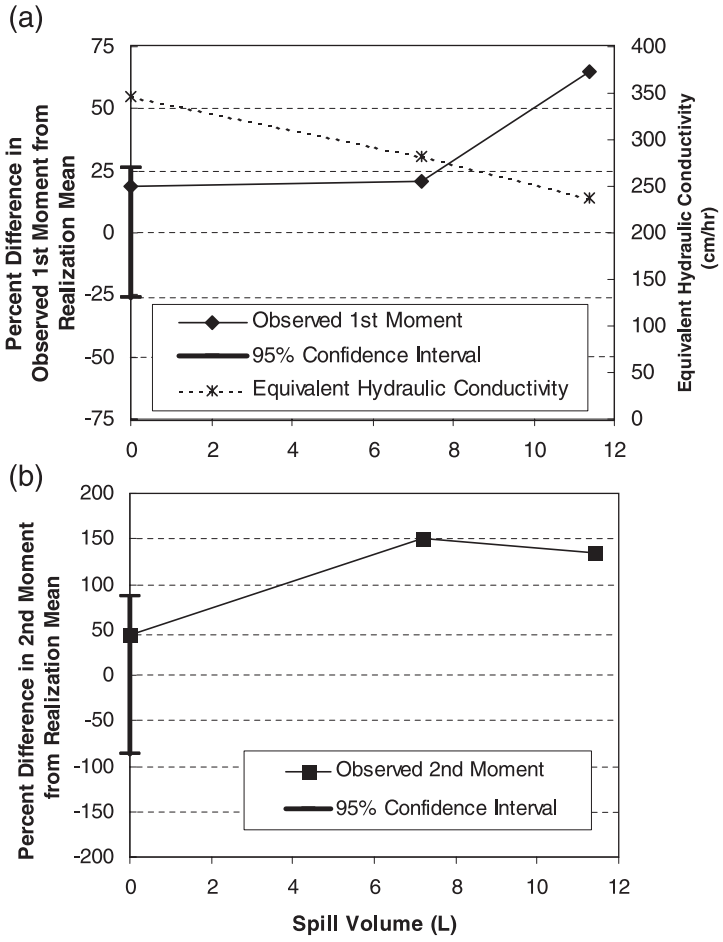


Fig. 8. Moments of the observed pre- and post-spill BTCs and the 95% confidence intervals based on simulations using 150 realizations of the pre-spill heterogeneous distribution and experiment C7 boundary conditions. The confidence intervals are indicated on the left axis corresponding to a zero volume spill. (a) First moments and equivalent hydraulic conductivity. (b) Second moments. The first moment is not affected by the first spill but changes significantly for the second spill exceeding the 95% confidence interval. The second moment changes significantly for the first spill and remains essentially unchanged for the second spill.

erably following the first spill. However, the second spill had limited additional impact on the second moment.

5. Numerical simulations

The finite-difference ground-water flow model, MODFLOW (McDonald and Harbaugh, 1988), was used to simulate pressure and flow in the tank. The details of the finite-difference grid representation are presented in Barth et al. (2001b). Simultaneous inverse modeling of the four pre-spill experiments was performed by nonlinear regression using MODFLOWP (Hill, 1992) to determine the values of K that simulated the observed head and flow data (K_r) (Barth et al., 2001b). The K_r values were used for all flow and transport simulations. Using the K_r values, head and flow predictions agreed with the observations. The variograms of $\ln K_r$ for the discretized distribution are shown in Fig. 6.

Transport was simulated using MT3DMS (Zheng, 1998) which uses the flows generated by MODFLOW or MODFLOWP. The third order, Total Variation Diminishing (TVD) solver was used. Forward transport simulations were performed using the K_r values. Tracer simulations using K_r values produced BTCs that match the observed pre-spill values quite well. Fig. 9 shows the predicted and observed BTCs from transect 2 during experiment C7. The other pre-spill experiments produced similar results. Injection of the tracer was represented as an initial concentration in two adjacent columns of finite-difference cells. This corresponded to the height of the injection interval and the width of the source immediately following injection. Additional details of the transport simulations are presented in Barth et al. (2001a,b).

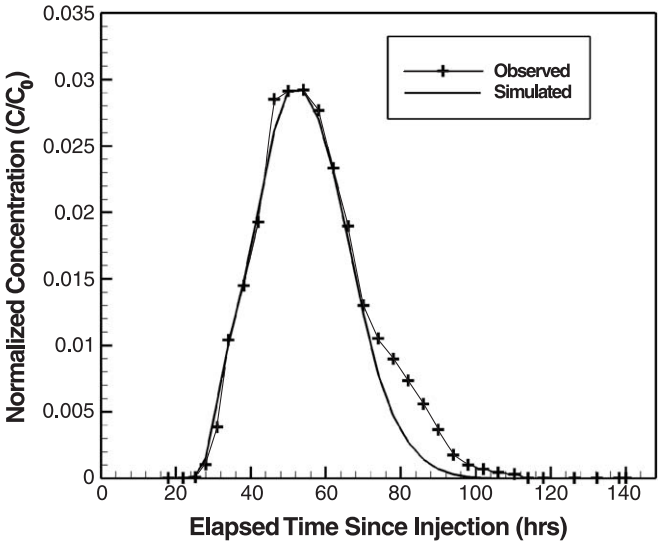


Fig. 9. Demonstration of calibrated conservative-transport model: observed and predicted BTCs for transect 2 prior to NAPL spill (experiment C7).

5.1. Multiple pre-spill realizations

There is no doubt that the presence of NAPL will have some impact on conservative-solute transport. The issue is whether that impact is significant. Does the presence of NAPL cause BTCs from tracer tests traversing a spill zone to be substantially different than BTCs from tracer tests in a clean portion of the same site? When NAPL is not present, differences in the exact spatial arrangement of subsurface materials between any two locations will result in different BTCs for different locations. However, statistical similarity can be expected if the distribution of ln$K$ in the aquifer is stationary and the tracer traverses a sufficient distance. In such cases, tracer tests from different locations effectively sample different realizations of a given distribution. To determine if a NAPL spill can have a significant impact, it is necessary to first determine typical tracer-test BTC variability that occurs when tracer tests are performed in different locations, or for different realizations, of an aquifer having a stationary distribution of ln$K$. This was accomplished by simulating pre-spill transport in 150 realizations of ln$K$ with the same statistical structure as the ln$K$ distribution packed in the tank. Boundary conditions from C7 were used for all realizations. The BTCs from each transect for each realization were evaluated and the results summarized in terms of the temporal moments. Statistics of the moments from the 150 realizations were used as an indication of the ensemble response of the heterogeneous distribution.

The gray BTCs in Fig. 7 represent BTCs from the 150 realizations of the heterogeneous distribution using K_r values and C7 boundary conditions. The figure also shows the ensemble average BTC: the average BTC across all the realizations. Interpretation of the NAPL-spill impact on the BTCs is qualitative; however, it does appear that at least the second spill had an impact that was significant compared to variability between realizations. To evaluate significance of the spills' impact, Fig. 8 includes the 95% confidence intervals of the simulated, pre-spill BTC moments. As expected, the observed and simulated results for the realization packed in the tank fall well within the confidence intervals. The observed BTC moments confirm that m_1 does not change significantly following the first spill: the value is within the pre-spill 95% confidence interval. However, the value of μ_2 following the first spill is outside the corresponding pre-spill 95% confidence interval. The second NAPL spill had a significant impact on m_1, shifting it well beyond the range of realization variability, while μ_2 hardly changed between the first and second spills (Fig. 8).

6. Discussion/conclusions

Detailed measurements and mapping provided a unique demonstration of the distribution of NAPLs in the saturated subsurface. One of the most dramatic features is the propensity for pore-scale features to control the NAPL distribution and the resulting pattern of sharply contrasted NAPL saturation in adjacent lenses. The mapping of NAPL saturation distribution, with explicit resolution of the heterogeneity, in an intermediate-scale heterogeneous porous medium provides a unique data set quantifying the behavior of NAPL. A review of the literature did not reveal previous experiments that have produced

such a data set. Such experimental data on NAPL saturations are potentially useful for evaluating models of NAPL migration in heterogeneous porous media, serving as a high-quality data set for model testing. In particular, it appears that models based on the continuum multiphase flow equations are not capable of predicting the pore-scale fingering, which seems to have caused many of the features noted in the experiment. The data could be used to test more recently developed percolation-theory models (e.g., Ewing and Berkowitz, 1998; Glass et al., 2001).

As should be expected, results from the experiments demonstrate that the presence of NAPL affects flow and transport. The details of the NAPL distribution, the affect on flow and transport, and the significance of these affects are the most interesting aspects of this work. For example, the effective K, calculated from the overall flow rate and gradient, consistently decreased with increasing amounts of entrapped NAPL. Transport observations, however, demonstrated that even in this two-dimensional system there is significant potential for bypassing: it was possible for much of the conservative tracer to bypass the first NAPL spill without causing a significant change in m_1. However, the second moment of conservative-tracer breakthrough was more sensitive to NAPL entrapment, even for smaller spill volumes.

Evaluation of the NAPL's impact on the heterogeneous distribution using the gamma-scan data corroborated the tracer-test data. Changes in the observed μ_2 from the tracer BTCs correspond quite well with the differences between pre- and post-spill variograms (Fig. 6): the first spill causes a large increase in $\sigma^2_{\ln K}$, which is reflected in μ_2, while the second spill causes only a small additional increase in $\sigma^2_{\ln K}$, and a corresponding minor change in μ_2.

The intermediate-scale demonstration of conservative-tracer sensitivity to macroscale entrapped NAPL encourages evaluation of the method's potential for field site applications. The basic objective would be the same: determine the circumstances for which a NAPL spill begins to significantly impact the temporal moments of a conservative tracer BTC. Basic constraints on the two-dimensional experiments would apply to the field site as well. For example, aquifer heterogeneity at the field site must be statistically stationary over a large enough region, as is the case for the Borden and Cape Cod sites. However, a fundamental difference between the two-dimensional experiments and the field would be the third dimension. In three dimensions there is added potential for bypassing macroscale entrapped NAPL. To offset the increased potential for bypassing, field applications will require an approach designed to maximize BTC sensitivity to the presence of entrapped NAPL. Tracer test parameters that may improve sensitivity include: delivery methods, timing, and the spatial distribution of injectors and samplers. For example, a forced gradient tracer test may be an effective way to reduce bypassing in a three-dimensional flow system. For any design, field-testing would require experiments with tracer tests before and after spilling NAPL, and from a pristine portion of the same site. Evaluating results from such a combination of design variations and field testing would provide insight on the potential of using conservative tracer tests to detect the presence of macroscale entrapped NAPL in the field, and is a topic of future research.

The results of these experiments, demonstrating macroscale entrapment in a heterogeneous porous medium and its impact on conservative-tracer transport, have implications for the application of partitioning tracer tests. The partitioning tracer test analysis

discussed in the literature (Jin et al., 1995) is based on the assumption that NAPL is present at residual saturation throughout the contaminated medium. Bypassing of flow around regions of high NAPL saturation, or even around individual NAPL-saturated lenses, will reduce contact between the partitioning tracer and the entrapped NAPL, reducing the impact of NAPL on partitioning-tracer retardation. The experimental results presented here suggest that when macroscale entrapment occurs, partitioning tracer test results will reflect either (1) the estimate of residual NAPL in a region that bypasses the NAPL spill, or (2) the invalid use of a conservative tracer BTC whose first moment has been significantly impacted by the presence of macroscale entrapped NAPL. Conservative tracer tests indicating a considerable amount of bypassing and macroscale entrapped NAPL could be used to modify the interpretation of partitioning tracer tests to more accurately reflect the conditions. Assessing whether sufficient bypassing has reduced or eliminated the efficacy of the method is a topic of future research.

The combination of size and complexity, in a laboratory setting, provided a unique opportunity to examine two consecutive NAPL spills in a heterogeneous porous medium, and the impact of the NAPL on transport. Comparing the NAPL's impact on tracer BTCs to variability of simulated BTCs from different realizations of the heterogeneous distribution allowed significance to be gauged and develop the following conclusions.

(1) There is a spill-size threshold below which there is no effect on the first moment: the first spill allowed considerable bypassing of the spill region as a whole and had no significant effect on the first moment. The second spill virtually eliminated the possibility of bypassing and had a dramatic impact on the first moment. For spills large enough to eliminate most bypassing, the first moment provides an indicator of entrapped NAPL.
(2) The second temporal moment of conservative transport BTCs is more sensitive to changes in the distribution of materials than the first moment. The first spill, which only occupied about 4% of the total pore volume, increased the second temporal moment significantly.

Acknowledgements

Financial support was provided by the US Geological Survey, US EPA Great Plains/Rocky Mountain Hazardous Substance Research Center, the US Army Research Office at Research Triangle Park, North Carolina and through National Science Foundation award EAR-0107095. Comments from Tarek Saba, Russell Detwiler, Steffen Mehl, two anonymous reviewers and Emil Frind helped to improve the manuscript are greatly appreciated.

References

Adams, E.E., Gelhar, L.W., 1992. Field study of dispersion in a heterogeneous aquifer, 2, spatial moments analysis. Water Resour. Res. 28 (12), 3293–3307.

American Society for Testing and Materials, 1994. Annual Book of ASTM Standards. Volume 04.08.

Barth, G.R., 1999. Intermediate-scale tracer tests in heterogeneous porous media: investigation of density, predictability and characterization of NAPL entrapment, PhD Dissertation. University of Colorado, Department of Civil Engineering. 360 pp.

Barth, G.R., Illangasekare, T.H., Hill, M.C., Rajaram, H., 2001a. Analysis of intermediate-scale tracer experiments for the development of tracer density guidelines. Water Resour. Res. 37 (1), 21–31.

Barth, G.R., Hill, M.C., Illangasekare, T.H., Rajaram, H., 2001b. Predictive and regression modeling of intermediate-scale flow and transport experiments in heterogeneous porous media. Water Resour. Res. 37 (10), 2503–2512.

Boggs, J.M., Young, S.C., Benton, D.J., Chung, Y.C., 1990. Hydrogeologic characterization of the MADE site, EPRI EN-6915, research project 2485-5, Interim Report. Electric Power Research Institute, Palo Alto, CA.

Brooks, R.H., Corey, A.T., 1966. Properties of porous media affecting fluid flow. ASCE, J. Irrig. Drain. Div., IR2 92, 61–88.

Compos, R., Illangasekare, T.H., Rajaram, H., 1998. Multiple experimental realizations of dense nonaqueous phase liquid spreading in saturated heterogeneous porous media, MS thesis. Dept. of Civil and Environmental Engineering, University of Colorado. 126 pp.

Dagan, G., Indelmann, P., Moltyaner, G., 1997. Stochastic analysis of concentration measurements in the transport experiment at Twin Lake Site. Water Resour. Res. 33 (4), 559–567.

Dai, D., Barranco, F.T., Illangasekare, T.H., 2001. Partitioning and interfacial tracers for differentiating NAPL entrapment configuration: column-scale investigation. Environ. Sci. Technol. 35 (24), 4894–4899.

Ewing, R.P., Berkowitz, B., 1998. A generalized growth model for simulating initial migration of dense nonaqueous phase liquids. Water Resour. Res. 34 (4), 611–622.

Freyberg, D.L., 1986. A natural gradient experiment on solute transport in a sand aquifer, 2, spatial moments and the advection and dispersion of non-reactive tracers. Water Resour. Res. 22 (13), 2031–2046.

Garabedian, S.P., LeBlanc, D.R., Gelhar, L.W., Celia, M.A., 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts 2. Analysis of spatial moments for a non-reactive tracer. Water Resour. Res. 27 (5), 911–924.

Glass, R.J., Conrad, S.H., Yarrington, L., 2001. Gravity-destabilized nonwetting phase invasion in macroheterogeneous porous media: near-pore-scale macro modified invasion percolation simulation of experiments. Water Resour. Res. 37 (5), 1197–1207.

Hess, K.M., Wolf, S.H., Celia, M.A., 1992. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts 3. Hydraulic conductivity variability and calculated macrodispersivities. Water Resour. Res. 28 (8), 2011–2027.

Hill, M.C., 1992. A computer program (MODFLOWP) for estimating parameters of a transient, three-dimensional groundwater flow model using nonlinear regression. U.S. Geological Survey Open File Report 91-484. 358 pp.

Illangasekare, T.H., Znidarcic, D., Walser, G., Weaver, J., 1994. An experimental evaluation of two sharp front models for vadose zone nonaqueous phase liquid transport, EPA/600/R-94/197, Robert S. Kerr Environmental Laboratory, Ada, OK.

Illangasekare, T.H., Armbruster, E.J., Yates, D.N., 1995a. Non-aqueous-phase fluids in heterogeneous aquifers, experimental study. J. Environ. Eng. 121 (8), 571–579.

Illangasekare, T.H., Ramsey, J.L., Jensen, K.H., Butts, M.B., 1995b. Experimental study of movement and distribution of dense organic contaminants in heterogeneous aquifers. J. Contam. Hydrol. 20, 1–25.

James, A.I., Graham, W.D., Hatfield, K., Rao, P.S.C., Annable, M.D., 2000. Estimation of spatially variable residual nonaqueous phase liquid saturations in nonuniform flow fields using partitioning tracer data. Water Resour. Res. 36 (4), 999–1012.

Jin, M., Delshad, M., Dwarakanath, V., McKinney, D.C., Pope, G.A., Sepekrnoori, K., Tilburg, C.E., 1995. Partitioning tracer test for the detection, estimation, and remediation performance assessment of subsurface nonaqueous phase liquids. Water Resour. Res. 31 (5), 1201–1211.

Kueper, B.H., Frind, E.O., 1991. Two-phase flow in heterogeneous porous media, 1, model development. Water Resour. Res. 27 (6), 1049–1057.

Kueper, B.H., Gerhard, J.I., 1995. Variability of point source infiltration rates for two-phase flow in heterogeneous porous media. Water Resour. Res. 31 (12), 2971–2980.

LeBlanc, D.R., Garabedian, S.P., Hess, K.M., Gelhar, L.W., Quadri, R.D., Stollenwerk, K.G., Wood, W.W., 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts 1. Experimental design and observed tracer movement. Water Resour. Res. 27 (5), 895–910.

Mackay, D.M., Freyburg, D.L., Roberts, P.V., Cherry, J.A., 1986. A natural gradient experiment on solute transport in a sand aquifer, 1, approach and overview of plume movement. Water Resour. Res. 22 (13), 2017–2029.

Mapa, J., Garcia, J., Illangasekare, T.H., 1994. Upscaling of water and solute transport in saturated porous media. Draft of Progress Report for U.S. Army Engineers Waterways Experiment Station.

Mayer, A.S., Miller, C.T., 1992. The influence of porous medium characterisitics and measurement scale on pore scale distributions of residual non-aqueous phase liquids. J. Contam. Hydrol. 11, 189.

McDonald, M.G., Harbaugh, A.W., 1988. A modular three-dimensional finite-difference ground-water flow model: U.S. Geological Survey Techniques of Water Resources Investigations. Book 6, Chap. A1. 586 pp.

Moltyaner, G.L., Wills, C.A., 1991. Local- and plume-scale dispersion in the Twin lake 40- and 260-m natural tracer tests. Water Resour. Res. 27, 2007–2026.

Rubin, Y., Ezzedine, S., 1997. The travel times of solutes at the Cape Cod tracer experiment: data analysis, modeling and structural parameters inference. Water Resour. Res. 33 (7), 1537–1547.

Shinozuka, M., Jan, C.-M., 1972. Digital simulation of random processes and its applications. J. Sound Vib. 25 (1), 111–128.

Sudicky, E.A., 1986. A natural gradient experiment on solute transport in a sand aquifer: spatial variability of hydraulic conductivity and its role in the dispersion process. Water Resour. Res. 22 (13), 2069–2082.

Szlag, D., 1997. Bioremediation of aquifer materials contaminated with organic chemicals, PhD dissertation. University of Colorado, Boulder, CO.

Walser, G.S., 1995. Vadose zone infiltration, mobilization and retention of nonaqueous phase liquids, PhD dissertation. University of Colorado at Boulder. 255 pp.

Woodbury, A.D., Sudicky, E.A., 1991. The geostatistical characterization of the Borden aquifer. Water Resour. Res. 27 (4), 533–546.

Zheng, C.M., 1998. A modular three-dimensional multispecies transport (MT3DMS) model. Release Do-D_3.00.A, prepared for the Waterways Experiment Station, U.S. Army Corps of Engineers.