Determination of DNAPL entrapment architecture using experimentally validated numerical codes and inverse modeling

S. Saenton and T.H. Illangasekare

aCenter for Experimental Study of Subsurface Environmental Processes (CESEP), Division of Environmental Science and Engineering, Colorado School of Mines, Golden, Colorado 80401, U.S.A.

The presence of dense non-aqueous phase liquid (DNAPL) in source zones in aquifers generates continuous mass flux long after the initial spill. It is our hypothesis that observed dissolved concentration in a monitoring well downstream of the source zone will provide inadequate and often misleading information on the entrapment architecture needed to design effective remediation schemes. An inverse modeling study was conducted to evaluate what additional information is needed to determine entrapment architecture. Three synthetic entrapment architectures generated by a multi-phase flow model UTCHEM in a spatially correlated random field were used in the analysis. The three selected from eighty realizations of the random field represent conditions that produce DNAPL pools, zones of residual DNAPL saturation and a combination of pools and residuals, respectively. For each of these three entrapment architectures, a data set of downstream concentration and solute mass flux was generated using a laboratory validated dissolution model based on MODFLOW and RT3D. An inverse modeling algorithm (PEST) was used to back-calculate the saturation distribution of DNAPL in the source zone. The inverse solution (i.e. DNAPL saturation distribution defining the architecture) was found to be non-unique when only concentration data was used. However when mass flux data that combines concentration and water flow, was added as an additional observation, the inverse problem converged rapidly to a unique solution. Further analysis to determine optimal monitoring strategies showed that the mass flux-matching technique could be used to determine entrapment architecture with some limitations. The technique becomes less accurate in terms of both total mass estimation and the ability to resolve the vertical distribution of DNAPL, when a source zone contains more pools than residuals. It was also found that with larger number of observations in the vertical direction (multi-level sampling points), the predictions become more accurate. The distance the monitoring well was placed downstream of the source zone affected the accuracy of prediction, but the estimate of the total mass entrapment was not affected. This is suspected to be an artifact of the two-dimensional test system that was used in this hypothetical analysis. More rigorous analysis using realistic three-dimensional systems is needed to make more definitive conclusions.

*Funding sources: National Science Foundation Awards EAR-0107095 and DMS-0222286, and Department of Defense SERDP Grant CU-1294.*
1. INTRODUCTION

Non-aqueous phase liquids (NAPLs) such as gasoline and chlorinated solvents are common organic compounds found at waste sites [1]. These organic liquids in the entrapment zone slowly dissolve into the flowing ground water generating a downstream contaminant plume in which concentrations are usually found to exceed regulatory standards. Mass transfer limitations due to rate-limited dissolution result in a long-term persistence of the source even in cases of small entrapped DNAPL mass. The identification of DNAPL sources located beneath the water table is critical to achieve goals of site remediation and aquifer restoration. Traditional soil coring is not widely used in DNAPL contaminated sites because it could remobilize entrapped DNAPLs [2]. In addition, the estimation of entrapped mass using point data obtained from cores is strongly sensitive to the choice of interpolation techniques and the estimates tend to be associated with high uncertainty [3].

The emergence of partitioning interwell tracer tests offers an in situ, non-intrusive site characterization method that minimizes the risk of remobilization of DNAPL [4]. This technique utilizes a suite of conservative and several partitioning tracers to detect the presence of DNAPL and to estimate the hydrodynamically accessible DNAPL mass in the entrapment zone [5]. Results obtained from this technique provide only the average saturation and total DNAPL volume in the tracer-swept zone. In addition, analyses of the tracer data are limited when using the method of temporal moments because equilibrium partitioning is assumed. This may not be the case when higher saturation zones of DNAPLs such as pools are encountered because the limited contact time between flowing aqueous phase containing tracers and DNAPL is not sufficient for equilibrium partitioning to take place. As a result, DNAPL volume is always underestimated [6,7].

Current research by Moreno-Barbero and Illangasekare [8] is attempting to improve the accuracy of this technique in detecting and distinguishing DNAPL in pools and intermediate to low saturation zones as well as estimating total NAPL volume. The methods that have been developed treat a pool as having two zones: a zone of high saturation at the bottom bounded at the top by a transition zone where NAPL saturation decreases gradually to residual. Preliminary results of this ongoing study indicate that multiple tracer tests detect the portion of DNAPL which is mostly in the transition zone whereas DNAPL in the high saturation zone is by-passed due to permeability reduction. This observation can be extended to dissolution that requires the flowing water to get in contact with the DNAPL. As a large fraction of DNAPL mass may reside in the high saturation portion of the pool, a smaller mass fraction in the transition zone contributes to a majority of mass flux. Knowledge of hotspots within the entrapment zone where the most DNAPL mass flux is produced is useful for remediation design and source zone management. Most of the removable DNAPL mass resides in the transition zone, and this region is more easily accessed by the injected treating reagents (e.g., surfactants, oxidants, and augmented microbes) during remediation, compared to the high saturation zone in the pool. In this study, we will investigate the feasibility and effectiveness of using observed steady-state concentration and dissolved DNAPL mass flux distributions downstream of the source in combination, to determine the entrapment architecture. An experimentally validated dissolution model in combination with an inverse modeling algorithm are used
2. BACKGROUND

2.1. Previous work

The use of pore water concentration distribution to infer the presence of DNAPL has been reported. Feenstra et al. [9] found that if pore water concentration of the dissolved contaminant is near the solubility limit (or effective solubility in case of multi-component DNAPL), DNAPL is presumed to be located within the vicinity of the sampling location. Anderson et al. [10] analyzed downgradient concentration patterns resulting from a well-defined (i.e. hypothetical) DNAPL source using a 3-D analytical solution and equilibrium mass transfer assumption. They found that residual DNAPL in ganglia contributes to high mass loadings during earlier phase of dissolution whereas pools, dissolving more slowly, remain as a source for a much longer time. Based on this observation, they proposed that the dissolved concentration (as a measure of dissolution rate) can be used to differentiate residual ganglia from pools. Although their findings can be used to qualitatively assess the relative importance of dissolved mass contributed by fingers and pools, the total DNAPL mass as well as actual entrapment architecture still cannot be determined.

Butcher and Gauthier [11] used an inverse analytical solution technique developed by Ala and Domenico [12] to determine residual entrapped DNAPL mass from measured mass flux emanating from the source zone. Instead of using an equilibrium dissolution model, both residual and pool non-equilibrium mass transfer models developed by Miller et al. [13] and Johnson and Pankow [14], respectively, were applied. The use of an analytical solution has several limitations; for example, the source zone architecture has to be simplified and ground water velocity is assumed uniform despite soil heterogeneity that causes complex flow patterns. Therefore the inverse modeling that was used resulted in a very high uncertainty in mass estimates. This study was an initial attempt to determine source zone architecture based on mass flux. However, the use of analytical solutions inevitably exclude the effect of inherent soil heterogeneity which always complicates the dissolution pattern and mass flux generation. In addition, the mixing due to advection and dispersion, and flow by-passing due to permeability contrast were neglected in this analysis. A simple source zone geometry must be assumed when analytical solutions are used conduct such analysis. In order to overcome these limitations, a comprehensive three-dimensional numerical model for mass transfer and solute transport is necessary for more realistic analysis at heterogeneous field sites with complex entrapment architecture.

2.2. Natural dissolution modeling

The dissolution of NAPL is usually modeled using a linear-driving force mass transfer expression, Eq. (1), described by:

\[ J_v = -k_{La} (c - c^*) \]

where \( J_v \) is a volumetric mass flux due to dissolution of NAPL per unit volume of porous medium, \( c \) and \( c^* \) are dissolved concentration and aqueous solubility, respectively. \( k_{La} \) is a lumped mass transfer coefficient which can be estimated from a phenomenological based
The empirical model such as Eq. (2) [15]:

\[ k_{La} = 12.41 (Re)^{0.227} (Sc)^{0.5} \left( \frac{\theta_n d_{50}}{\tau L^*} \right)^{1.276} \left( \frac{D_m}{d_{50}^2} \right), \]  

(2)

where Re is Reynolds number (representing hydrodynamic condition of the system), Sc is Schmidt number (chemical dependent), \( \theta_n \) is volumetric NAPL content, \( d_{50} \) is mean particle diameter, \( \tau \) is tortuosity factor, \( L^* \) is the characteristic or dissolution length, and \( D_m \) is molecular diffusion coefficient. The dissolution module that was validated in the laboratory using a two-dimensional dissolution cell [15] is incorporated into the existing RT3D code in the RCT or reaction package [16]. Transient mass transfer is simulated using a series of steady-state simulations. With the velocity fields obtained from MODFLOW [17], RT3D together with a dissolution module simulate mass transfer for a given porosity and NAPL saturation. The simulation sequence is repeated until the target time is reached, updating values of NAPL saturation and effective hydraulic conductivity at the end of each simulation time step. Such automated sequential simulation using separate models of flow, dissolution and transport is made possible using PERL language programming that allows for the transfer of output from one simulation to code to another. Detailed discussion on model development and implementation can be found in Saenton [15].

3. NUMERICAL SIMULATION EXPERIMENTS OF CONCENTRATION AND FLUX MATCHING

3.1. 2-D test tank

A set of numerical simulations was conducted to represent conditions in a two-dimensional, intermediate-scale, heterogeneously-packed hypothetical test tank with dimensions of 9.53-m long, 1.02-m high, and 0.05-m thick. A test tank was used in this hypothetical example with the goal of conducting future experiments in a laboratory test system to generate data to validate methods that are presented in this paper. The tank consists of homogeneous and heterogeneous zones (see Fig. 1). Two constant-head tanks at both ends are used to maintain a steady hydraulic gradient throughout the simulation. The homogeneous part of the tank consists of #8 sand within which a fully-penetrated injection well is placed to deliver tracers. The heterogeneous zone is packed using five test sands of different sieve sizes (#16, #30, #50, #70, and #110) representing hydraulic conductivity values in a range of 3 orders of magnitude. The heterogeneity was designed as a spatially correlated random field with statistical parameters similar to field sites reported in literature [18]. This heterogeneous packing assumes a log-normal distribution of hydraulic conductivity with \( \mu_{lnK} = 4.18 \) (K is in cm hr\(^{-1}\)) and a variance \( \sigma^2_{lnK} \) of 1.22. The correlation lengths in the lateral (\( \lambda_h \)) and vertical (\( \lambda_v \)) directions are 0.508 and 0.0508 m, respectively. The heterogeneous zone consists of 1280 cells of 0.254 m in length and 0.0254 m in depth which corresponds to 32 columns and 40 layers resulting in 16 lateral and 20 vertical correlation lengths.

3.2. DNAPL source zone

The tetrachloroethene or PCE, was used as the test DNAPL in this study. This is the same DNAPL that was used in developing the dissolution model given in Eq. (2). The
DNAPL entrapment in the source zone (1.54-m×1.02-m) was created using UTCHEM to numerically simulate a spill. The location of the source zone is shown in Fig. 1. The two-dimensional source zone was discretized into 40 layers and 24 columns. The capillary pressure-saturation relationship needed in the multiphase flow simulation assumes Brooks-Corey model. The total PCE spill volume is 0.001 m$^3$ (the resultant average saturation $\bar{S}_n$ is 0.0359) and the spill rate is $1.93 \times 10^{-8}$ m$^3$ s$^{-1}$. The point of spill is located at the center of the top layer to represent a surface spill. The PCE spill simulation was conducted for a long enough time period for the PCE distribution to reach a static state.

Three selected examples of model simulated PCE distributions used in the analysis are shown in Fig. 1. As can be seen from the distribution pattern, PCE source zone consists of both high saturation pools and residuals. In realization #54, most of the spilled PCE forms pools on top of low permeability layers. On the other hand, spill #11 produced PCE saturations in the range of intermediate to low values. Realization #1 shows a combination between the previous two entrapment configurations where both pools and residuals are formed throughout the depth of the source zone.

### 3.3. Concentration/Flux matching method: an inverse modeling study

The concentration/flux matching method for determining the NAPL saturation profile through inverse modeling is presented. For brevity, this method will be referred to as C&FM technique. In general, in the inverse problem, an attempt is made to match the model simulated values with the measurements (called observations) by adjusting model input parameter(s). In this study, the model input parameter is NAPL saturation distribution, whereas the observations are the downstream concentrations and mass flux.

#### 3.3.1. procedures

The procedure used in inverse modeling to determine the three synthetically generated entrapment architectures is presented. For each of the selected architectures, a natural
dissolution simulation was conducted to generate the steady-state concentration and mass flux dataset that was used as observations. These are equivalent to field observations in monitoring wells. Another dissolution model was also setup, but with a re-conceptualized NAPL source zone as shown in Fig. 2. In this conceptual source zone, the NAPL saturation distribution was assumed homogeneous for the entire layer, but could vary from one layer to another. Then, the inverse modeling algorithm was used to systematically adjust these NAPL saturations such that the model simulated downstream concentrations and mass fluxes matched the observations. In addition to evaluating the C&FM technique for its applicability to quantify the NAPL distribution, the optimal monitoring strategies for placement and sampling at the observation wells were also studied. The monitoring strategies, in this context, refer to the sampling frequency or the number of observation points in the vertical directions (40, 20, and 10 evenly distributed observation points), and the sampling location or the distance downstream the monitoring well is installed (MW1, MW2, and MW3).

Figure 2. Synthetically real source zone used to obtain the observations vs. the re-conceptualized source zone of homogeneously distributed NAPL within a layer.

For the test case with 40 observation points, there were a total of 80 observations consisting of 40 concentration and 40 mass flux values. Thus, it is statistically plausible to determine the NAPL saturation for all 40 model layers through a non-linear regression since the number of observations is greater than the number of parameters. However, in the case for 20- and 10-observation points, it is statistically unreasonable to determine the NAPL saturation for all 40 layers. Instead, the model was re-discretized into 20 and 10 layers for both cases, respectively. In this way, only twenty and ten NAPL saturations were evaluated.

The C&FM technique is a parameter estimation process that requires an efficient and robust inverse modeling algorithm. Universal inverse modeling code PEST (Parameter ESTimation) developed by Doherty [19] was used instead of another popular code UCODE [20]. A major advantage of PEST over UCODE is that optimum parameter values can be constrained to lie between individually specified upper and lower bounds. The problem investigated in this study involves the estimation of NAPL saturation $S_n$ that has to be in the range of $S_n \in [0, 1]$. If the permissible upper and lower bounds of $S_n$ is not provided,
the possibility exists for the inverse algorithm to compute to physically unrealistic NAPL saturation values. As the dissolution module relies on realistic saturation values, the inverse modeling simulations will terminate prematurely.

Multiple tracer test simulations were also setup using the partitioning tracer test module developed by Saenton [15], for all three spills in order to determine the entrapment architecture, and to compare with the results obtained from inverse modeling. The multiple tracer test simulation involves the injection of a pulse of conservative and partitioning tracer solution into the test aquifer through a fully penetrating injection well (see Fig. 1). The downstream tracer concentrations were monitored at MW1 through multi-level sampling of each of the forty model layers. The observed breakthrough curve for each layer (i.e., sampling location) was analyzed using the method of temporal moments to obtain the average saturation in the layer. The NAPL saturations were then converted to total mass in each layer. This mass distribution profile was compared with both actual mass distribution and the results obtained from the C&FM technique.

Inverse modeling using the C&FM technique with only observed concentrations (i.e., no mass fluxes) produced non-unique solutions to the problem posed. In other words, there were more than one set of NAPL saturation profiles that resulted in the same downstream concentration. However, when the mass fluxes were added as additional observations (i.e., constraints), the inverse problem converged rapidly to a unique solution. Therefore, both concentrations and mass fluxes were used in all subsequent analyses.

4. RESULTS AND DISCUSSION

This section presents the results obtained from the numerical study to determine DNAPL entrapment architecture using the multiple tracer test and the C&FM technique. Results from both techniques will also be compared with the actual mass distribution in order to assess their accuracy. The effects of entrapment architecture, sampling frequency, and sampling locations on the inverse modeling results will be addressed.

4.1. Effect of entrapment architecture on multiple tracers test and C&FM technique

Figs. 3, 4, and 5 show the vertical mass distribution profiles obtained from both the multiple tracer test and the C&FM technique. These results are based on measurements at MW1 with the 40-observation-point multi-level sampler. Results from both techniques are also compared with the actual mass distributions for all three realizations. Table 1 shows total PCE mass estimated from both techniques. In all cases, the flux matching technique provided good estimates of NAPL distribution (in the vertical direction), whereas the mass distribution profile from the tracer technique were not accurate, especially when the PCE source zone contained mostly high saturation blocks (e.g., spill #54). In terms of total mass estimation, both techniques performed equally well, although a large fraction of PCE in the high saturation zone remained undetected (see results from realization #54 as an example). The more residuals the source contains, the better the estimates of total mass and mass distribution were likely to be. In these examples, total PCE mass was better estimated for entrapment architecture #11 (mostly low saturation zones) than #1. The estimates for #11 and #1 were both better than #54 (mostly high saturation zones).

It was also found that total mass flux emanating from the entrapment zone strongly
depended on the amount of tracer detectable mass. Fig. 6 shows the relationship between total mass flux emanating from the source zone due to natural dissolution as a function of tracer detectable PCE mass, or alternatively, the hydrodynamically accessible mass. Based on these simulations, it can be concluded that the higher the hydrodynamically accessible mass in the source zone (e.g., spill #11), the higher the observed the downstream dissolved mass flux. These examples show that the mass flux distribution not only can be used to infer source zone architecture, the total mass flux can be used to estimate the total hydrodynamically accessible mass. The entrapped DNAPL mass in the hydrodynamically accessible zones is the primary contributor to the dissolved mass in the downstream plume.

Figure 3. PCE mass distribution: actual mass distribution, and results from multiple tracers test and inverse modeling for spill #1 (40 observation points).

Figure 4. PCE mass distribution: actual mass distribution, and results from multiple tracers test and inverse modeling for spill #11 (40 observation points).

Table 1
Mass estimates from multiple tracers test and C&FM technique. The results are based on 40-observation-point multi-level sampling well MW1.

<table>
<thead>
<tr>
<th>Realization</th>
<th>#1</th>
<th>#11</th>
<th>#54</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total PCE mass (kg)</td>
<td>1.62</td>
<td>1.62</td>
<td>1.62</td>
</tr>
<tr>
<td>PCE mass from tracer test (%)</td>
<td>82.0</td>
<td>94.1</td>
<td>20.3</td>
</tr>
<tr>
<td>PCE mass from inverse modeling (%)</td>
<td>68.2</td>
<td>89.0</td>
<td>55.7</td>
</tr>
</tbody>
</table>
4.2. Effect of sampling frequency on C&FM technique

As shown in the previous section, vertical profiles of PCE mass based on the flux matching technique for all three cases were relatively accurate. These results were not surprising as a dense sampling array (40 pts.) was used. The sampling captured all the mass fluxes emanating from the source, thus providing accurate information for inverse modeling. However, if the number of sampling points is smaller, it is expected that the estimated mass distribution profile as well as total mass may not be accurate.

Fig. 7 shows mass distribution profiles for all three cases of multilevel sampling that used 40-, 20-, and 10-layer observations monitored at MW1 respectively for spill #1. Realization #1 was selected in this discussion as its entrapment architecture contains both high saturation pools as well as residuals. From these plots, it can be seen that the accuracy of the predicted mass distribution decreased with the number of sampling points. Furthermore, the predicted PCE mass in the source zone also decreased with the reduction of the number of observations (see Table 2). From the results of these examples, it can be concluded that the prediction accuracy of both mass distribution profile and total mass improve with increasing number of observations.

4.3. Effect of sampling locations on C&FM technique

The accuracy of estimated mass as a function of observation distance downstream of the source zone was evaluated. Three monitoring well locations (MW1, MW2, and MW3) were used to collect information on concentrations and mass fluxes. This information was then used as observations in the C&FM technique. Fig. 8 shows PCE mass distribution profiles estimated from inverse modeling for 20 observation points at MW1, MW2, and MW3 respectively for spill #1. These results demonstrate that when observations were taken at farther distance downstream of the entrapment zone, the predicted mass distribution...
profiles became less accurate. The vertical concentration distribution at longer distances downstream is more uniform due to mixing by advection and dispersion controlled by the soil heterogeneity. Consequently, the use of a less informative concentration profile resulted in a less accurate mass distribution profile.

Table 3 shows the total estimated percent PCE mass using 20-point-observation scheme for realization #1. Interestingly, the estimates of total PCE mass for all three cases were similar despite the differences in mass distribution profiles. This in a limited way demonstrates the benefits of using the mass flux distribution data in addition to concentrations for inverse modeling.
5. CONCLUSIONS

In this theoretical study, we demonstrated through a set of numerical examples that a concentration and flux matching method can be used to determine hydrodynamically accessible DNAPL volume as well as entrapment architecture in the source zone. Although it is a computationally intensive method, it works equally well as costly multiple tracers technique (or even better in case of a pool-containing source zone), and can be considered an alternate choice for site characterization. As found in any site characterization technique, this method has some limitations. The technique becomes less accurate in terms of both total mass estimation and the ability to resolve the vertical distribution of DNAPL when a source zone contains more pools than residuals. It was also found that with a larger number of observations in the vertical direction (multi-level sampling points), the predictions become more accurate. The distance the monitoring well was placed downstream of the source zone affected the accuracy of prediction, but the estimate of the total mass entrapment was not affected.

Although this technique was tested in a two-dimensional domain, this method should also be applicable to a real 3-D system or a field-scale problem. This is because the basic processes such as flow by-passing that affect mass transfer in heterogeneous systems are reasonably captured in a two-dimensional flow field. Nevertheless, this newly developed concentration/flux matching technique has yet to be verified using data from experiments or field tests prior to a real implementation.
Table 3
Percent of actual PCE mass estimated from inverse modeling using 20-point observation at MW1, MW2, and MW3 for spill #1.

<table>
<thead>
<tr>
<th>Monitoring locations</th>
<th>PCE mass estimates (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MW1</td>
<td>61.4</td>
</tr>
<tr>
<td>MW2</td>
<td>61.1</td>
</tr>
<tr>
<td>MW3</td>
<td>59.8</td>
</tr>
</tbody>
</table>

REFERENCES