Numerical modeling and data-worth analysis for characterizing the architecture and dissolution rates of a multicomponent DNAPL source

Andres E. Prieto-Estrada[1], Mark A. Widdowson[1], and Lloyd D. Stewart[2]

[1] The Charles E. Via, Jr. Department of Civil and Environmental Engineering, Virginia Tech, Blacksburg, Virginia 24061-0105, United States
[2] Praxis Environmental Technologies, Inc., 1440 Rollins Road, Burlingame, California 94010, United States

Corresponding author, e-mail address: mwiddows@vt.edu (M.A. Widdowson)

Key Points

- Aqueous-phase concentrations monitored in a field experiment were simulated to quantify NAPL distribution and dissolution rates
- Depletion profiles of the most soluble DNAPL component accurately constrained the source zone architecture
- Multiscale heterogeneity of source zone architecture controlled the uncertainty of estimated mass transfer coefficients

Abstract

A numerical solute transport model was history matched to a high-resolution monitoring dataset to characterize a multicomponent source of nonaqueous phase liquids (NAPLs) and evaluate the uncertainty of estimated parameters. The dissolution of NAPL mass was simulated using the SEAM3D solute transport model with spatially-varying NAPL saturations and mass transfer rate coefficients, representing the heterogenous architecture of the source zone. Source zone parameters were simultaneously estimated using PEST from aqueous-phase concentrations measured in a multilevel monitoring transect and from mass recovery rates measured at extraction wells during a controlled field experiment. Data-worth analyses, facilitated by PEST ancillary software, linked maximum aqueous-phase concentrations of all compounds to reductions in prior uncertainty of mass transfer coefficients. In turn, transient concentrations of the most soluble NAPL fraction constrained the source mass estimation. Accurately estimating the source mass and reducing prior uncertainties was possible by removing concentrations measured during early NAPL dissolution stages, identified as prior-data conflicts using the iterative ensemble smoother PESTPP-iES. Prior-based Monte Carlo analyses highlighted model limitations for representing sub-grid-scale heterogeneity of source zone architecture and NAPL dissolution, yet history-matching of final dissolution stages measured at multilevel ports eliminated parameter bias and produced long-term projections of source depletion with multistage behavior. Including mass discharge constraints further improved the accuracy of source mass estimation, complementing multilevel monitoring constraints on the source architecture and mass transfer coefficients.

Plain Language Summary

Quantifying dissipation timeframes of DNAPL source zones and contaminant discharge rates is fundamental for environmental-management support. Both variables depend on the distribution of DNAPL mass (architecture) within the source zone, and cannot be quantified by direct observation methods. We elucidated upon the worth of multilevel monitoring for characterizing the source zone architecture of a field experiment with inverse numerical modeling of contaminant transport and DNAPL dissolution. Uncertainties on estimated DNAPL distribution and dissolution rates were primarily associated to variability in dissolved concentration trends at multiple scales. Dissolved concentration peaks measured during early DNAPL dissolution stages were found responsible for inducing model parameter and predictive errors. Yet the depleting signature of the least soluble component accurately constrained the source zone architecture, combining mass recovery rates with multilevel monitoring to reduce model uncertainties. Hence, our approach and results have beneficial implications for management support of aged source zones undergoing final depletion stages.

Index Terms and Keywords

1831 Groundwater Quality, 1846 Model calibration (3333), 1873 Uncertainty quantification (3275), 4314 Mathematical and computer modeling, 1815 Monitoring, forecasting, prediction (4315)

DNAPL source zone, numerical modeling, uncertainty quantification, DNAPL mass transfer, source zone architecture, contaminant mass discharge

1. Introduction

Remediation and long-term dissipation of contaminant source zones comprised of dense nonaqueous phase liquids (DNAPLs) in the subsurface encompasses technical challenges related to uncertainty of DNAPL spatial distribution and dissolution rates (Kueper et al., 2014; Mayer & Hassanizadeh, 2005; NRC, 2005). Entrapped DNAPL mass and saturation distributions in the porous medium, referred to as the source zone "architecture", are key parameters controlling source-zone longevity and depletion behavior (Dekker and Abriola, 2000; DiFilippo & Brusseau, 2008). Typical multistage and nonmonotonic depletion profiles observed in monitoring data reflect the gradual dissolution of NAPL accumulations with characteristic saturations (Brusseau et al., 2013; Kokkinaki et al., 2014; Stewart et al., 2021). Accumulations of low-saturation ganglia allowing for groundwater flow-through account for peaks of discharge concentrations at early NAPL dissolution stages, whereas high-saturation pools with negligible hydraulic accessibility account for dissolution tailing at the final stages of a source lifespan (Christ et al., 2010; Kueper et al., 2014; Yang et al. 2018). Thus, estimating the depletion behavior and remedial timeframes of a source zone requires knowledge on NAPL architecture, which is difficult to characterize with direct observation methods (Engelmann et al., 2019), but can be estimated from field tests and monitoring data using mathematical models of NAPL dissolution (Falta et al., 2005a, 2005b; Stewart et al., 2021).

Numerical modeling methods coupling groundwater flow and contaminant transport with NAPL dissolution have been used to estimate saturation distributions and mass transfer rate coefficients from monitoring data (Frind et al., 1999; Mobile et al., 2012; Saenton & illangasekare, 2004). Researchers have focused on estimating NAPL architecture or depletion timeframes from synthetically-generated source zones and aqueous-phase concentrations using several mathematical approaches to simulating mass transfer. Several studies considered either a local equilibrium assumption (LEA) or Gilland-Sherwood models of interphase mass transfer (Kang et al., 2021a, 2021b; Marble et al., 2008; Saenton & Illangasekare, 2004). The applicability of LEA in decision-support models is questionable because heterogeneity of aquifer hydraulic properties and DNAPL architecture can induce flow bypassing and mass transfer rate limitations, resulting in nonequilibrium concentrations typically observed at field sites (Falta, 2003; Kokkinaki et al., 2013; Powers et al., 1992, 1994). Similarly, Gilland-Sherwood models rely on correlations between empirical coefficients and soil particle sizes that were determined under specific bench-scale conditions, which may not be applicable to field-scale problems with different hydraulic conditions (Powers et al., 1994; Saenton & Illangasekare, 2007). Additional uncertainties on LEA and Gilland-Sherwood models include grid discretization requirements, as both approaches have been validated with pore-scale experimental data (Agaoglu et al., 2015; Falta, 2003).

Upscaled models have been developed to simulate NAPL dissolution kinetics over a representative elementary volume (REV) incorporating source zone metrics (Christ et al., 2010; Marble et al., 2008; Parker & Park, 2004; Saenton & Illangasekare, 2007; Stewart et al., 2021; Zhu & Skyes, 2004). These metrics include NAPL mass and descriptions of source zone architecture in the form of areal dimensions of NAPL accumulations or the ganglia-to-pool (GTP) mass ratio metric (Abriola et al., 2013; DiFilippo & Brusseau, 2011). Because these metrics are difficult to measure at contaminated sites, upscaled models incorporating a spatially-varying lumped-process mass transfer coefficient have also been used to interpret monitoring data and predict source depletion timeframes (Guo et al., 2020; Marble et al., 2008; Mobile et al., 2012; Park & Parker, 2005). These models simplify the heterogeneity of porous media, aqueous-phase velocities, NAPL architecture, and dispersivity, into a single lumped-process parameter at the REV scale (Falta, 2003; Imhoff et al., 1993; Luciano et al., 2018). Although scale-dependent mass transfer rate coefficients may simplify grid discretization requirements, the parameterization of NAPL source zones for inverse numerical modeling and uncertainty quantification with spatially-correlated random parameter fields is not straightforward (Arshadi et al., 2020; Kang et al., 2021a, 2021b; Kock & Nowak, 2015, 2016)

Given that NAPL source zones have complex spatial morphologies with sharp saturation transitions at fine scales, traditional interpolation and geostatistical methods used in groundwater flow modeling may be not suitable for parameterizing NAPL source zones (Arshadi et al., 2020; Kang et al., 2021a). Alternative methods proposed for parameterizing NAPL source zones include deep learning algorithms trained with images of saturation distributions generated with multiphase flow simulations on highly-resolved permeability fields (Arshadi et al., 2020; Kang et al., 2021a, 2021b), posing additional data requirements and uncertainties on porous media characteristics and model parameters (Abriola, 1989; Agaoglu et al., 2015; Miller et al., 1998). Moreover, these parameterization methods have been tested with synthetically-generated source zones to estimate categories of NAPL saturations through inverse modeling conditioned by borehole data (Arshadi et al., 2020), or by aqueous-phase concentrations under LEA (Kang et al., 2021a, 2021b). Although

these methods can generate physically-based, spatially-correlated categorical parameters, they are computationally expensive and require further validation and verification with field data.

Numerical models with parsimoniously parameterized source zones have proved useful for characterizing NAPL architecture and/or lumped-process mass transfer coefficients (Marble et al., 2008; Mobile et al., 2012; Saneton & Illangasekare, 2004). Moreover, combining multilevel monitoring with recovery rates of contaminant mass (or with conventional monitoring wells) can be valuable for characterizing heterogeneous NAPL architectures, as spatially-varying contaminant fluxes may be mapped to soil horizons harboring NAPL mass within a source zone (McMillan et al., 2018). Several studies have incorporated Gilland-Sherwood or upscaled mass transfer functions in discretized NAPL zones or in dual-domain models to estimate grid-scale parameters from multilevel monitoring data and/or mass discharge/flux measurements (Christ et al., 2010; Falta, 2003; Frind et al., 1999; Guo et al., 2020; Mobile et al., 2012; Park & Parker 2005; Saenton & Illangasekare, 2004). Although previous investigations have demonstrated the utility of parameterization parsimony for characterizing NAPL source architecture and dissolution rates with inverse modeling, the uncertainty of grid-scale parameters arising from the assimilation of high-resolution monitoring data has not been investigated.

A primary objective in this numerical modeling study was to evaluate the worth of aqueous-phase concentrations monitored at a multilevel sampling (MLS) transect in combination with total mass discharge (MD) rates to quantify and reduce the uncertainty of the mass, architecture, and dissolution rates of a multicomponent DNAPL source. The field experiment considered in this study involved the creation a small-scale source zone with a heterogeneous architecture at the Borden experimental site (Broholm et al., 1999). Broholm et al. (2005) quantified the initial mass of the DNAPL mixture using multicomponent NAPL dissolution theory and mass balance analyses. Mobile et al. (2012) estimated the source architecture and dissolution rates from MD profiles and from a single, incomplete MLS nest, using an inverse modeling technique. They constrained the initial NAPL mass by the known amount and the source zone dimensions by the measured post-experimental footprint (Mobile et al., 2012). Our study expanded on both previous analyses by incorporating the entire MLS transect to infer the source footprint and to quantify initial NAPL mass, saturation distribution, and mass-transfer rate coefficients, while examining causality of parameter uncertainty. A secondary objective was to investigate model limitations for reproducing the observed system behavior, further elucidating upon monitoring data assimilation for source zone characterization with inverse numerical modeling, while minimizing the propensity for biasing timeframes of source zone persistence.

2. Materials and Methods

2.1. Overview of Field Experiment of Multicomponent DNAPL Dissolution

This study incorporated a dataset documenting the dissolution of a DNAPL mixture in a field experiment (Broholm et al., 1999). At the Borden experimental site in Canada, a 5-liter (7.7 kg) mixture of dyed solvents (10% TCM, 40% TCE, and 50% PCE by volume) was injected at approximately 0.05 m below the water table into a 55 m^3 (5.5 m long x 4.5 m wide x 2.3 m deep) unconfined aquifer test cell (Figure 1) comprised of medium- to fine-grained lacustrine sand with occasional beds of coarse sand/gravel and silt. Groundwater flow through the test cell, bounded

laterally by sheet piling on four sides, was maintained by a network of five upgradient injection and five downgradient extraction wells. The network was operated at approximately 360 L/d, generating a mean groundwater velocity of 0.13 m/d and horizontal hydraulic gradient of 0.023 ± 0.0024. Aqueous-phase contaminant concentrations were monitored for 220 days in a fence of MLS ports (Figure 1) located approximately 2.1 m downgradient from the DNAPL release location. After 220 days of natural dissolution monitoring, a 5.5-day pulse of methanol was injected in the test cell to evaluate dissolution enhancement processes (Broholm et al., 1999; Broholm, 2006). The horizontal spacing of MLS ports was 0.5 m with a vertical spacing of 0.1 m, with screen lengths less than 1 cm.

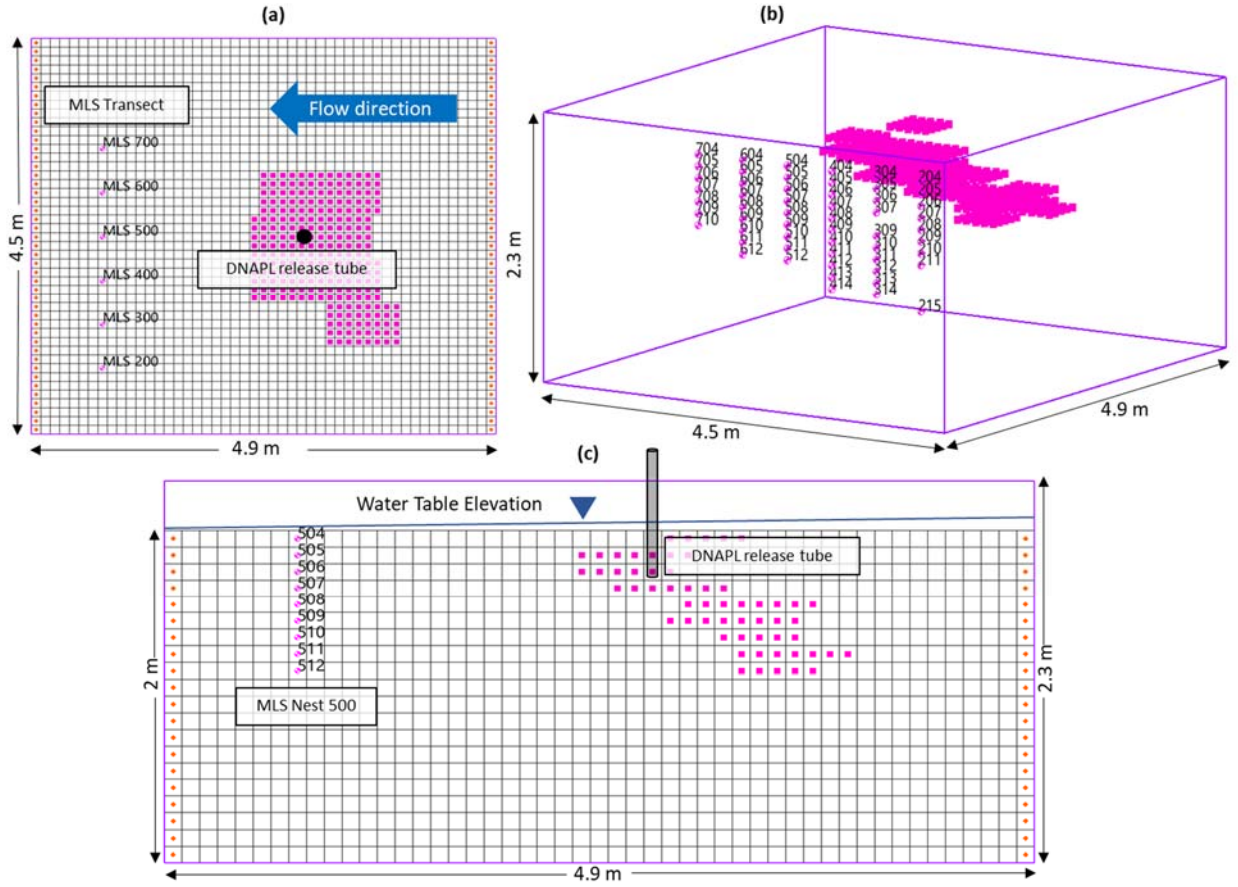

Figure 1. Configuration of aquifer test cell and its representation in the numerical model. (a) Plan view of all MLS nests and NAPL grid blocks encompassing the entire source zone footprint in the numerical model. (b) Test cell

Flow-weighted concentrations were monitored at the extraction wells with screen lengths spanning the average height (1.82 m) of the saturated zone for 291 days. The cell was excavated 291 days after the mixture release to map the DNAPL distribution over 5-cm vertical intervals (Figures S1 and S2) (Broholm et al., 1999). Broholm et al. (2005) compared estimates of initial NAPL mass calculated from the post-excavation source footprint and by multicomponent NAPL dissolution theory with mass balance analysis of effluent data and area- and depth-integrated MLS concentrations. Their best mass estimates ranged from 6.7 to 7.5 kg by averaging of MLS data, and differences between theoretical NAPL dissolution rates and effluent data were attributed to pre-flushing TCM dissolution and volatilization losses. Also, dissolved concentrations below

equilibrium levels observed throughout the experiment were attributed to dilution effects (Broholm et al., 2005). These results support kinetic mass transfer modeling coupled with uncertainty analyses for indirect characterization of NAPL source zones, as detailed, pore-scale process modeling and site characterization in this context is unfeasible for remedial-decision support at hazardous waste sites.

2.2. Numerical Modeling of Groundwater Flow and Contaminant Transport

A steady-state groundwater flow model was developed with MODFLOW2000 (Harbaugh et al., 2000). The aquifer test cell was discretized in 23 layers with uniform grid blocks measuring 10 cm along all dimensions. As shown in Figure 1, the extent of the model domain matched the size of the test cell along the vertical dimension (Z-axis) and the horizontal dimension perpendicular to the flow direction (Y-axis). The horizontal dimension parallel to flow (X-axis) was reduced from 5.5 m (test cell length) to 4.9 m for computational efficiency. Constant values of hydraulic conductivity, flow boundary conditions, and transport parameters were assigned to match field conditions, which were characterized with tracer tests and soil cores by Broholm et al. (1999) and analyzed with numerical modeling in Mobile et al. (2012). Model layers 1 through 3 were inactivated because the water table fell below the elevation of layer 3 during the monitoring period. Model layer 4 encompassed MLS port 504, which showed evidence of NAPL presence at 0.4 m below the top of the aquifer test cell despite the reported average depth of the water table during the experiment at ~0.48 m (Broholm et al., 1999).

Dissolution of the multicomponent DNAPL source and aqueous-phase contaminant transport were simulated with SEAM3D (Waddill & Widdowson, 2000). Interphase mass transfer [M T^{-1} L^{-3}] was simulated using a linear driving force model:

$$J_i = k_i^N (C_i^{eq} - C_i) \tag{1}$$

where k_i^N [T^{-1}]is a lumped mass transfer rate coefficient specific to each NAPL phase constituent i, C_i^{eq} [M L^{-3}] is the equilibrium solubility calculated according to Raoult's Law, and C_i [M L^{-3}] is the aqueous phase concentration. Equation 1 is coupled in SEAM3D to the following relationship representing NAPL dissolution from the soil medium into the aqueous phase:

$$J_i = -\rho_b \frac{dC_i^N}{dt} \tag{2}$$

where ρ_b [M L^{-3}] is the bulk density of the soil and C_i^N [M M^{-1}] is the NAPL mass of compound i per unit mass of dry soil. A modified version of SEAM3D incorporates the upscaled NAPL dissolution model developed by Parker and Park (2004) to simulate transient mass transfer rates:

$$k_i^N = k_{i,0}^N \left(\frac{\bar{q}}{\bar{K}}\right)^{\alpha} \left(\frac{M_i(t)}{M_{i,0}}\right)^{\beta} \tag{3}$$

where $k_{i,0}^N$ = initial mass transfer rate coefficient [T^{-1}], $\bar{q}$ = average Darcy velocity [L T^{-1}], $\bar{K}$ = average hydraulic conductivity [L T^{-1}], $M(t)/M_0$ = transient ratio of NAPL mass [M M^{-1}], α and β are dimensionless empirical parameters. Previous investigations have reported a linear relationship

between k_i^N and $\bar{q}$, with α = 1 (Parker & Park, 2004; Park & Parker, 2005). The transient mass ratio raised to the empirical depletion exponent β represents a reduction of NAPL/water interfacial areas over time, regulating tailing of discharge concentrations with reducing mass transfer rates as the source mass is depleted (Parker & Park, 2004; Stewart et al., 2020). A previous modeling study of the same Borden experiment indicated model insensitivity to the β exponent, attributed to a lack of extensive monitoring of decreasing discharge concentrations (Mobile et al., 2012). In this work, both the α and β parameters were set to zero to focus the uncertainty and data-worth analyses on the spatially-variable parameters $k_{i,0}^N$ and C_0^N. Adjusting C_0^N parameters allowed to estimate the initial mass (M_0^N [M]) of the entire NAPL mixture and quantify its uncertainty, as SEAM3D generates an output of remaining NAPL mass in the source zone every time step using Equation 4, by adding the mass of all grid blocks where a C_0^N parameter value was assigned:

$$M_0^N = \sum_{Zone\ 1}^{Zone\ 23} V^N \times C_0^N \times \rho_b \quad (4)$$

where V^N [L^3] = NAPL zone volume. The source zone was represented with 23 NAPL zones (Figures S1 and S2) positioned upgradient of MLS ports which showed contaminant breakthrough, suggesting the upgradient presence of NAPL mass (Figures 2, 3, and 4). Each NAPL zone was comprised of 25 grid blocks (Figures S1 and S2) and was assigned one pair of adjustable parameters, k_0^N and C_0^N, representing uniform mass distribution and dissolution within each NAPL zone. The areal dimensions of all NAPL zones were designed as 0.5 m x 0.5 m on the horizontal plane representing the horizontal spacing of MLS ports, whereas vertical layers of 0.1 m represented the vertical spacing between ports. The location of NAPL zones was determined by contaminant travel times analyzed from MLS breakthrough data. Overall, the source distribution in the model encompassed the observed post-excavation footprint (Figures S1 and S2), which likely developed through vertical and downward NAPL migration throughout the experiment (Broholm et al., 1999, 2005).

2.3. Parameter Estimation and Uncertainty Quantification

For each NAPL zone, C_0^N and k_0^N were simultaneously estimated from monitoring data. In addition, two global k_0^N multipliers to identify compound-specific mass transfer coefficients for TCE and PCE ($k_{TCE,0}^N$ and $k_{PCE,0}^N$) were estimated as multipliers of $k_{TCM,0}^N$ for a total of 48 adjustable source zone parameters. History-matching targets included 1,556 measurements of dissolved TCM, TCE, and PCE concentrations monitored at the MLS transect (Figure 1), out of 4,770 measurements comprising the entire MLS dataset. The 1,566 MLS targets corresponded to 23 ports including concentration measurements through 130 days (Figures 2, 3, and 4), when extraction well redevelopment abruptly increased the water table by 1 m causing significant data noise through 220 days (Broholm et al., 1999; Mobile et al., 2012). These MLS targets were grouped by sampling port and a weight of 1 was assigned to each aqueous-phase concentration measurement within each port. Additional constraints included 78 measurements of mass discharge rates monitored at extraction wells for 220 days (before methanol remediation was implemented). Mass discharge measurements were grouped by contaminant (i.e., three MD groups) and assigned uniform weights, balancing the initial error contribution of each group to the objective function (Φ). Specifically, individual measurements of MD were assigned a weight of 250 within the TCM and TCE MD groups, whereas individual PCE MD measurements were

assigned a weight value of 500. This weighting strategy was designed to balance the visibility of MD groups with MLS ports for Φ minimization.

An initial parameter estimation with PEST_HP (Doherty, 2020), which uses a parallelizable gradient-based optimization process, only included the MLS targets. All k_0^N and C_0^N parameters were log-transformed to facilitate the nonlinear optimization process. Bounds for C_0^N parameters were designed as a function of initial NAPL saturation (S_0^N [%]) of the pore space calculated as:

$$S_0^N = \frac{\rho_b C_0^N}{\rho_N \theta} \quad (5)$$

where ρ_N [M L^{-3}] is the NAPL density and θ is the soil porosity. Bounds for C_0^N parameters were set as $0.05 < S_0^N$ (%) < 25 for most NAPL zones, and as $0.05 < S_0^N$ (%) < 5 for NAPL zones below layer 10 and in layer 4. The prior (pre-history matching) NAPL mass value was set as ~120 % greater than the known initial mass (7.7 kg). This prior value was established to evaluate whether initial history-matching of MLS data with PEST_HP could result in a total NAPL mass estimate close to the known value. Bounds for all k_0^N ($k_{0,TCM}^N$) parameters were set as $0.01 < k_0^N$ (d^{-1}) < 7.5 following an order-of-magnitude range obtained through a simplified mass transfer correlation defined in Frind et al. (1999), where the dissolution of a large-scale DNAPL mixture in the Borden aquifer was simulated using a similar grid scale. Bounds for k_0^N compound-specific multipliers were kept consistent with ratios determined by Mobile et al. (2012) as $0.95 < k_{0,TCE}^N < 1$ and $0.8 < k_{0,PCE}^N < 0.95$.

The posterior uncertainty of NAPL mass was quantified with the iterative ensemble smoother PESTPP-iES (White et al., 2020). PESTPP-iES undertakes Monte-Carlo sampling of parameter uncertainty bounds generating ensembles which are upgraded with the Gauss-Levenberg-Marquardt (GLM) optimization algorithm. Rather than fitting simulation results to data, PESTPP-iES can generate observation ensembles considering a multi-gaussian distribution of measurement noise ($\boldsymbol{\varepsilon}$) (White, 2018). Here, σ_ε was defined as 5% of measured values. This stochastic approach was used for history-matching of (i) MLS data only, and (ii) both MLS and MD data, quantifying the posterior uncertainty of parameters (C_0^N, k_0^N) and predictions (M_0^N). In the following sections, Model A = optimized with PEST_HP using MLS data only, Model B = optimized with PESTPP-iES using MLS data only, and Model C = optimized with PESTPP-iES combining MLS and MD data.

Parameter bounds were used to define 95% confidence intervals of multi-gaussian prior probability distributions (PDF) of model parameters, assuming statistically-uncorrelated NAPL zones. The upper C_0^N bounds for NAPL zones were reduced from 25% to 15% S_0^N and initial parameter values were set from PEST_HP results. Prior to parameter upgrading, PESTPP-iES undertakes a prior-based Monte Carlo analysis to detect "prior-data conflicts" (PDC), which are measurements that cannot be simulated with the structural and parametrization design of the model (White et al., 2020, 2021). All PDCs flagged by PESTPP-iES were removed to eliminate history-matching induced bias, which would otherwise produce erroneous parameter values compensating for model defects (Doherty, 2015). Moreover, PESTPP-iES tracks the evolution of a "base realization" during the optimization process, corresponding to the initial parameter realization upgraded

without **ε** ensembles. Estimates of NAPL mass removed by natural dissolution produced by Model A were compared to estimates generated with the posterior base realization of Models B and C.

2.4. Data-Worth Analysis

First-order second-moment (FOSM) analysis was used to evaluate data worth for reducing the uncertainty of model parameters and predictions through history-matching. The GENLINPRED and PREDUNC utilities from the PEST software suite (Watermark Numerical Computing, 2018) were used for FOSM analyses. The primary assumption in FOSM analysis is model linearization expressed as:

$$\mathbf{h} = \mathbf{Zk} + \boldsymbol{\varepsilon} \tag{6}$$

which states that a vector of measurements of system state **h** (aqueous-phase concentrations) equals the action of the model **Z** (Jacobian sensitivity matrix weighted by $\sigma_{\varepsilon}^{-1}$) on a vector of parameters **k** plus a vector of measurement noise **ε** (Doherty, 2015). In this case, $\sigma_{\varepsilon}^{-1}$ was calculated on the basis of misfit between measurements and model outputs using the PEST-based utility PWTADJ2 (Watermark Numerical Computing, 2018) after history-matching. GENLINPRED and PREDUNC calculate the posterior uncertainty variance of model parameters through covariance propagation:

$$\mathrm{C}'(\mathbf{k}) = \mathrm{C}(\mathbf{k}) - \mathrm{C}(\mathbf{k})\mathbf{Z^t}[\mathbf{Z}\mathrm{C}(\mathbf{k})\mathbf{Z^t} + \mathrm{C}(\boldsymbol{\varepsilon})]^{-1}\mathbf{Z}\mathrm{C}(\mathbf{k}) \tag{7}$$

where the posterior covariance matrix C'(**k**) is obtained through history-matching (Doherty, 2015). In this case, the prior covariance matrix C(**k**) is diagonal with no spatial correlations between NAPL zones. The estimated initial NAPL mass, a SEAM3D output, was treated as a linearized model prediction:

$$s = \mathbf{y^t k} \tag{8}$$

where **y** is a vector of sensitivities of s with respect to **k**. The prior (σ_s^2) and posterior ($\sigma'_s{}^2$) uncertainty variances of s were calculated as:

$$\sigma_s^2 = \mathbf{y^t}\mathrm{C}(\mathbf{k})\mathbf{y} \tag{9}$$

$$\sigma'^2_s = \mathbf{y^t}\mathrm{C}'(\mathbf{k})\mathbf{y} \tag{10}$$

The worth of individual MLS ports for reducing prior parameter uncertainties was calculated with model A, whereas the worth of compound-specific MLS and MD datasets was calculated with the posterior base realization of model C. Model C was also used to quantify the worth of individual MLS and MD measurements and to quantify the relative uncertainty variance reduction (RUVR) of each parameter (i), defined as:

$$RUVR_i = 1 - \frac{\sigma'^2_i}{\sigma_i^2} \tag{11}$$

where σ_i^2 are prior parameter variances encapsulated in C(**k**) and σ'^2_i are posterior parameter variances extracted from C'(**k**) (Doherty, 2015).

3. Results and Discussion

3.1. Parameter Estimation and Uncertainty Quantification

A comparison of estimated NAPL mass and k_0^N multipliers is presented in Table 1. In general, the known NAPL mass (7.7 Kg) was underestimated (6.4 – 7.2 Kg) when history-matching to MLS data only, particularly with gradient-based optimization (PEST_HP). An improvement in the accuracy of mass estimation with PESTPP-iES was achieved through the removal of 206 PDC values flagged by PESTPP-iES (Figures 2, 3, and 4) comprising 13% of the MLS constraints. Likewise, 10 PDCs (Figure 5) representing 13% of the MD dataset were also flagged by PESTPP-iES and removed for history-matching. Including the MD constraints resulted in an excellent agreement with the known initial source mass, which was encompassed by 95% confidence limits (Table 1). Prior-based Monte Carlo results suggested that emphasizing early peak concentrations for history-matching can result in underestimation of NAPL mass and overestimation (bias) of k_0^N values, leading to underestimation of source dissipation timeframes. In this case, removing PDC values, rather than modifying the model design, was sufficient to accurately estimate NAPL mass and constrain mass transfer coefficients.

Table 1. Model-estimated DNAPL mass and k_0^N multipliers.

Parameter/Prediction	PEST_HP (A: MLS)	PESTPP-iES (B: MLS)	PESTPP-iES (C: MLS and MD)
Mass (Kg)	6.367	$\mu = 7.187$ $\sigma = 0.120$	$\mu = 7.626$ $\sigma = 0.110$
$k_{0,TCE}^N$ (d^{-1})	1.00	$\mu = 0.99$ $\sigma = 0.006$	$\mu = 0.99$ $\sigma = 0.005$
$k_{0,PCE}^N$ (d^{-1})	0.95	$\mu = 0.93$ $\sigma = 0.018$	$\mu = 0.88$ $\sigma = 0.017$

μ = mean estimated value. σ = standard deviation of estimated parameters (k_0^N) and predictions (Mass).

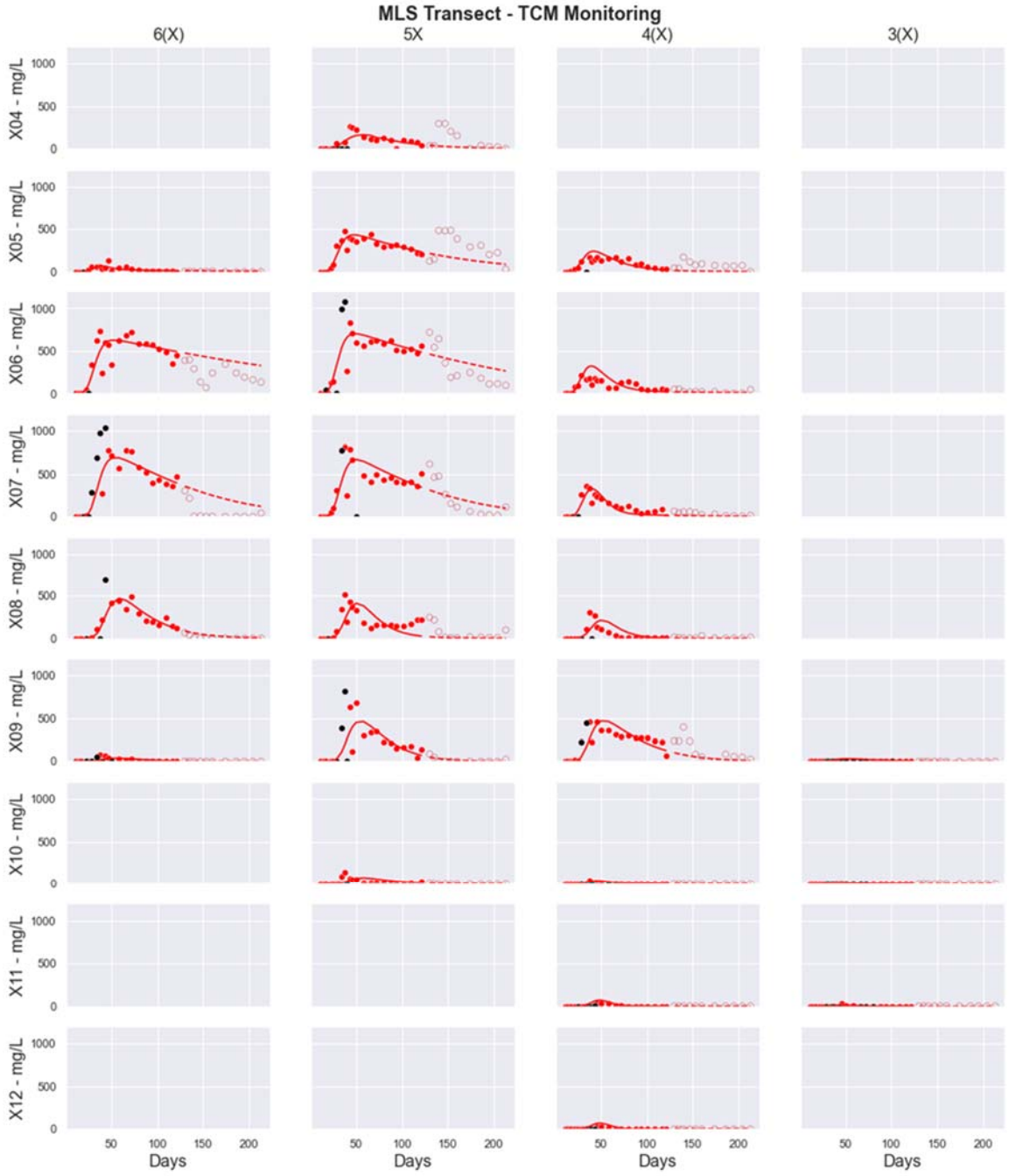

Figure 2. Comparison of measured (circles) and simulated (lines) aqueous-phase TCM concentrations at the MLS fence. Simulation results correspond to the posterior base realization of model C, including MLS and MD constraints. Empty circles correspond to concentrations ignored for history-matching because of significant measurement noise induced by water table fluctuations after 130 days of monitoring. The dashed lines correspond to simulated values beyond 130 days, informed by MD data exclusively. Black-filled circles are prior-data conflicts removed from the history-matching process to avoid parameter bias.

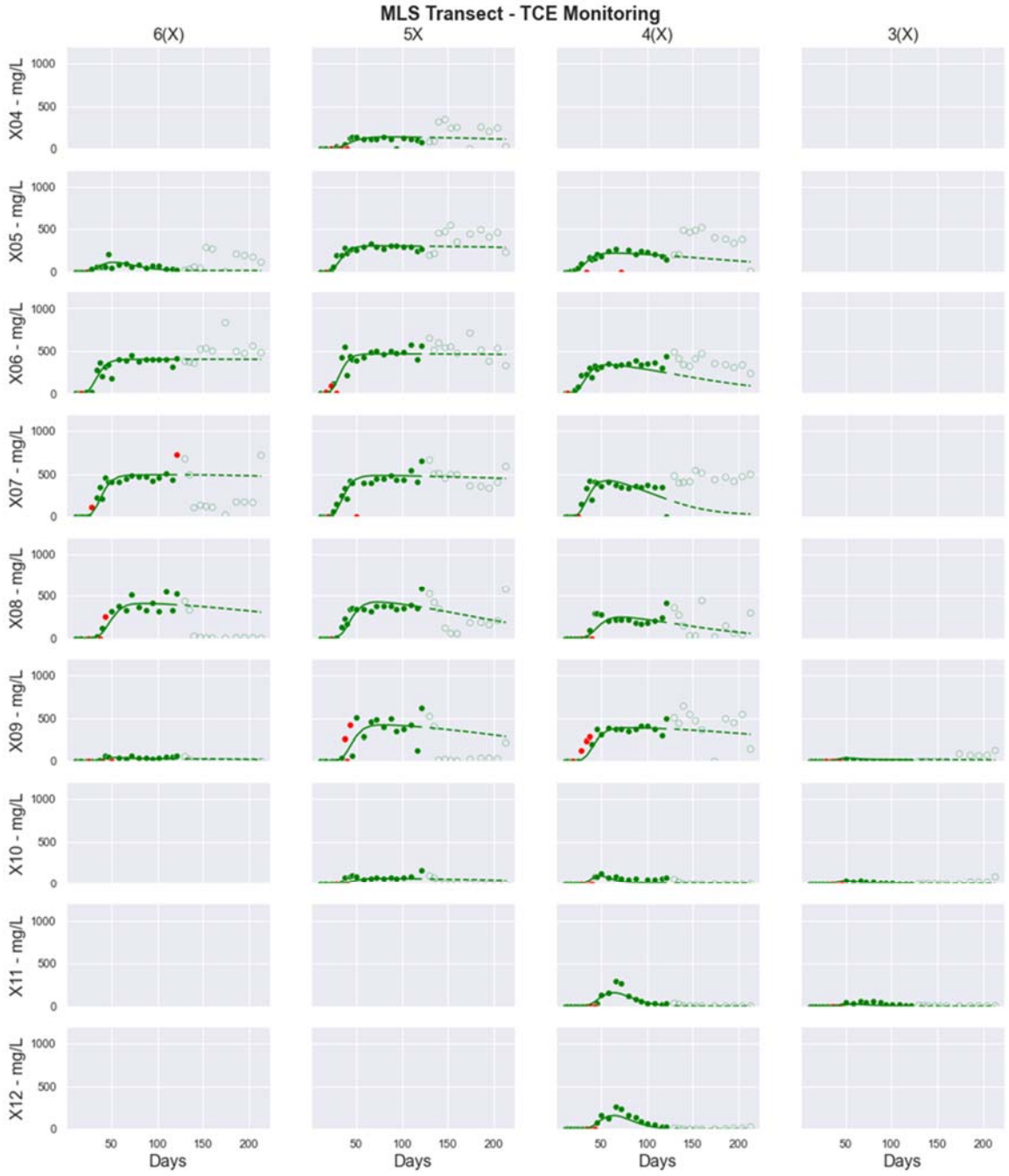

Figure 3. Comparison of measured (circles) and simulated (lines) aqueous-phase TCE concentrations at the MLS fence. Simulation results correspond to the posterior base realization of model C, including MLS and MD constraints. Empty circles correspond to concentrations ignored for history-matching because of significant measurement noise induced by water table fluctuations after 130 days of monitoring. The dashed lines correspond to simulated values beyond 130 days, informed by MD data exclusively. Red-filled circles are prior-data conflicts removed from the history-matching process to avoid parameter bias.

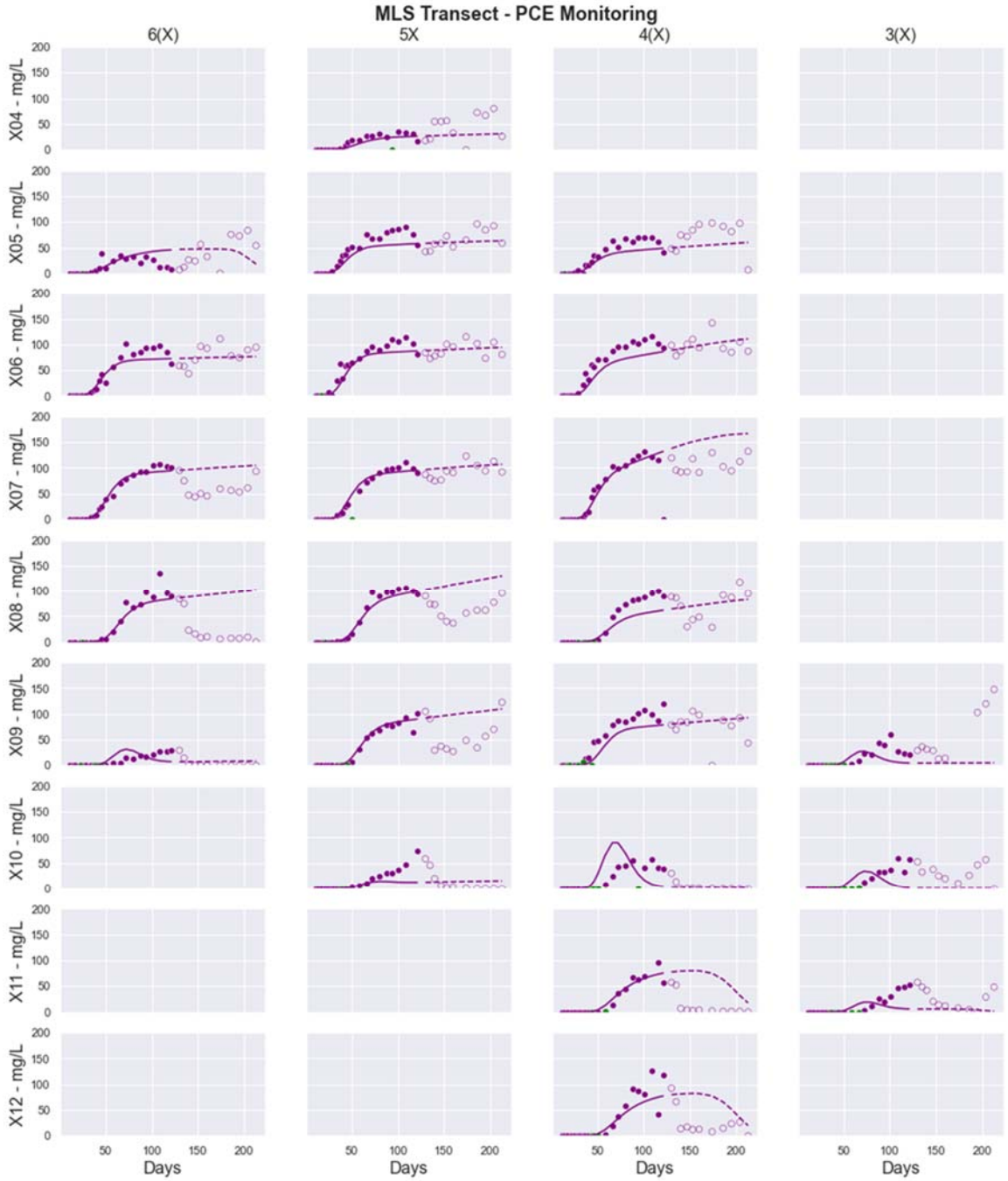

Figure 4. Comparison of measured (circles) and simulated (lines) aqueous-phase PCE concentrations at the MLS fence. Simulation results correspond to the posterior base realization of model C, including MLS and MD constraints. Empty circles correspond to concentrations ignored for history-matching because of significant measurement noise induced by water table fluctuations after 130 days of monitoring. The dashed lines correspond to simulated values beyond 130 days, informed by MD data exclusively. Green-filled circles are prior-data conflicts removed from the history-matching process to avoid parameter bias.

Prior-data conflicts pertaining to each dissolved NAPL component were detected at similar locations along the MLS and MD profiles (Figures 2 through 5). Most PDCs corresponded to initial TCM concentration peaks, some were detected along TCE breakthrough, and a few before PCE breakthrough. This may have been associated to propagation rates of component-specific mass

transfer zones in the NAPL as the source architecture developed. In theory, the dissolution process of NAPL mixtures has been described as a chromatographic process, where component-specific mass transfer zones propagate at different velocities through NAPL accumulations as a function of their local solubilities and their length along the principal flow direction (Geller & Hunt, 1993; Soerens et al., 1998). In practice, the numerical discretization of NAPL zones along the flow direction may influence the estimation of k_0^N values, as grid-scale concentration gradients (Equation 1) would also regulate the sequential mass transfer process for any prescribed q (Darcy velocity) and $\alpha_{L,T}$ (dispersivity) values (Falta, 2003; Frind et al., 1999; Hunt & Sitar, 1988). In this work, the unknown source zone architecture along the flow direction may have placed additional uncertainties on q, $\alpha_{L,T}$, and $k_{i,0}^N$ parameters, which regulate overall NAPL mass transfer at the grid scale (Rivett & Feenstra, 2005).

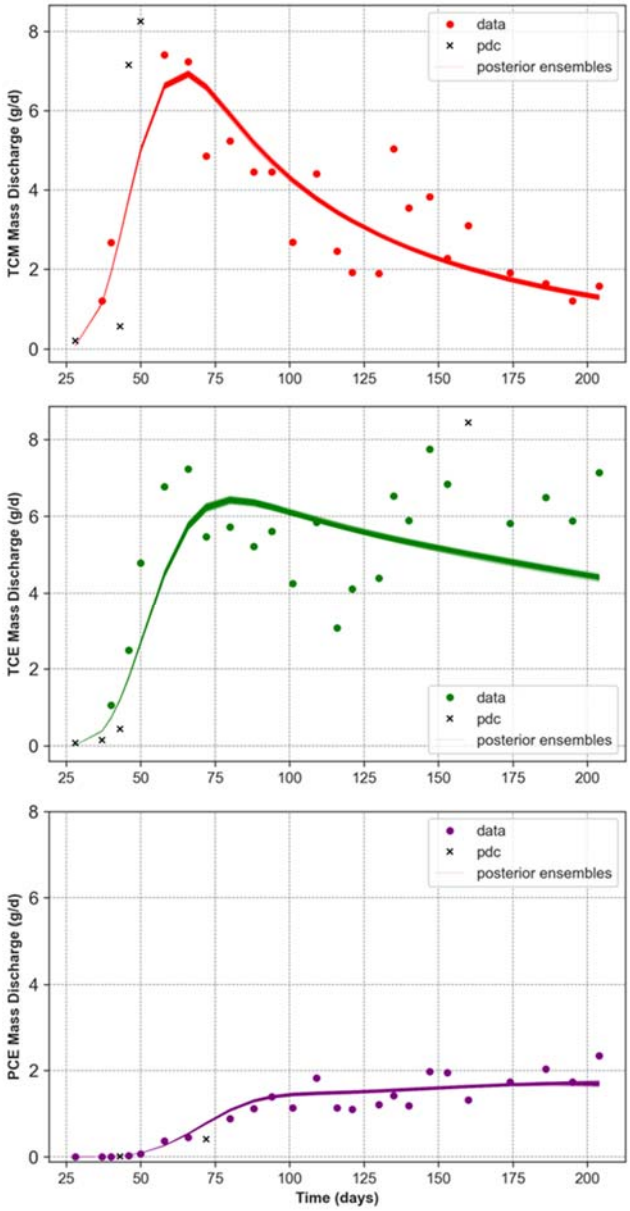

Figure 5. Posterior ensembles of MD profiles generated with model C.

Table 1 indicates that both $k_{i,0}^N$ multipliers estimated with Model A showed a tendency to reach their upper bounds. Only the values of $k_{PCE,0}^N$ estimated with PESTPP-iES were constrained within their prior uncertainty bounds (Table 1). In addition to the uncertain effects of grid scale on potential parameter bias, multistage NAPL dissolution below the MLS scale may have impacted posterior results. For example, inspecting the TCM signature of ports 506 and 507 in Figure 2 suggested two slopes of declining concentrations before 130 days. An initially steep slope between the concentration peaks through day 50, followed by a more gradual slope through day 130, suggested heterogeneity of NAPL architecture and dissolution below the MLS scale. Despite removing PDCs, sub-grid-scale multistage NAPL dissolution cannot be adequately simulated with a single parameter set (C_0^N, k_0^N) per NAPL zone. This explained why model A produced lower mass estimates with an increased propensity for biasing (overestimating) $k_{TCM,0}^N$ parameters (e.g., port 608 in Table 2), whereas model B (PDC targets removed) produced $k_{TCM,0}^N$ values consistent with model C (Table 2). Although these results suggested that a dual-domain approach may have better captured TCM profiles at MLS ports, removing PDCs to estimate a single-domain parameters accurately constrained NAPL mass and a consistent range of $k_{0,TCM}^N$ values.

Table 2. Distribution of estimated DNAPL mass and mass transfer coefficients

NAPL	MLS	$k_{TCM,0}^N$ (day^{-1})			NAPL Mass (Kg)		
Zone	Port	A	B	C	A	B	C
4	404	0.037	0.041	0.041	0.14	0.14	0.15
5.1	605	0.038	0.029	0.037	0.02	0.02	0.01
5.2	505	0.102	0.104	0.106	0.59	0.59	0.64
5.3	405	0.050	0.061	0.067	0.12	0.13	0.13
6.1	606	0.174	0.171	0.175	1.23	1.63	1.79
6.2	506	0.252	0.226	0.236	1.16	1.32	1.43
6.3	406	0.147	0.138	0.131	0.10	0.11	0.13
7.1	607	0.735	0.282	0.303	0.79	0.93	0.96
7.2	507	0.183	0.231	0.251	0.79	0.81	0.85
7.3	407	0.303	0.255	0.245	0.11	0.11	0.11
8.1	608	3.259	0.220	0.202	0.33	0.35	0.36
8.2	508	0.227	0.254	0.204	0.15	0.17	0.20
8.3	408	0.108	0.115	0.082	0.07	0.07	0.09
9.1	609	0.385	0.019	0.048	0.00	0.00	0.00
9.2	509	0.248	0.196	0.216	0.29	0.30	0.31
9.3	409	0.314	0.184	0.171	0.35	0.38	0.40
9.4	309	0.010	0.010	0.066	0.01	0.00	0.00
10.1	510	0.024	0.033	0.016	0.03	0.02	0.04
10.2	410	0.015	0.068	0.534	0.01	0.01	0.01
10.3	310	0.010	0.015	0.049	0.01	0.00	0.00
11.1	411	0.056	0.097	0.081	0.02	0.03	0.02
11.2	311	0.010	0.017	0.084	0.02	0.01	0.00
12	412	0.077	0.109	0.083	0.02	0.03	0.02
Total NAPL Mass (Kg)					6.4	7.2	7.6

A (PEST_HP, MLS only), B (PESTPP-iES, MLS only), C (PESTPP-iES, MLS and MD)

Long-term projections of MD using the posterior base realization of Model C (Table 2) are compared to the model calibrated by Mobile et al. (2012) in Figure 6. In both cases, multistage dissolution profiles emerged from the spatial distribution of NAPL mass and dissolution rates, as MD projections were produced with $\beta = 0$ (Equation 3). The mass transfer rates shown in Table 2 encompassed the same order-of-magnitude range reported in Mobile et al. (2012), although the values determined in this study were lower by ~50% on average, reflecting the grid-scale dependence of estimated k_0^N values. Specifically, the grid-block size in Mobile et al. (2012) of 500 cm^3 was also 50% smaller than the 1000 cm^3 scale used in this study. Furthermore, stochastic model optimizations with more adjustable NAPL zones than available MLS ports, using 5-cm thick layers, resulted in mass overestimation by orders of magnitude and inconsistent k_0^N and C_0^N distributions (results not shown).

3.2. FOSM-Analysis Results

The worth of monitoring datasets for reducing prior uncertainty variance of NAPL mass is shown in Figure 7. This figure highlights the importance of TCM data for constraining NAPL mass in contrast to the negligible worth in the monitoring profiles of other dissolved components. Likewise, Figure 7 shows the increase in data worth for reducing prior uncertainty of NAPL mass by removing PDCs. Although the short-term TCM MLS signatures (< 130 days) alone constrained the prior uncertainty of NAPL mass by ~100%, adding the complete TCM MD signature (220 days) improved mass estimates by spreading $\boldsymbol{\varepsilon}$ induced by water table fluctuations across all NAPL zones. Furthermore, Table 3 indicates a decrease in the worth of MLS ports commensurate with the S_0^N of their corresponding upgradient NAPL zones. This apparent correlation reflected the similarity between the TCM dissolution profile of each MLS port and the TCM MD profile, emphasizing the indirect value of multilevel monitoring for characterizing NAPL distribution and reducing the uncertainty of source depletion rates.

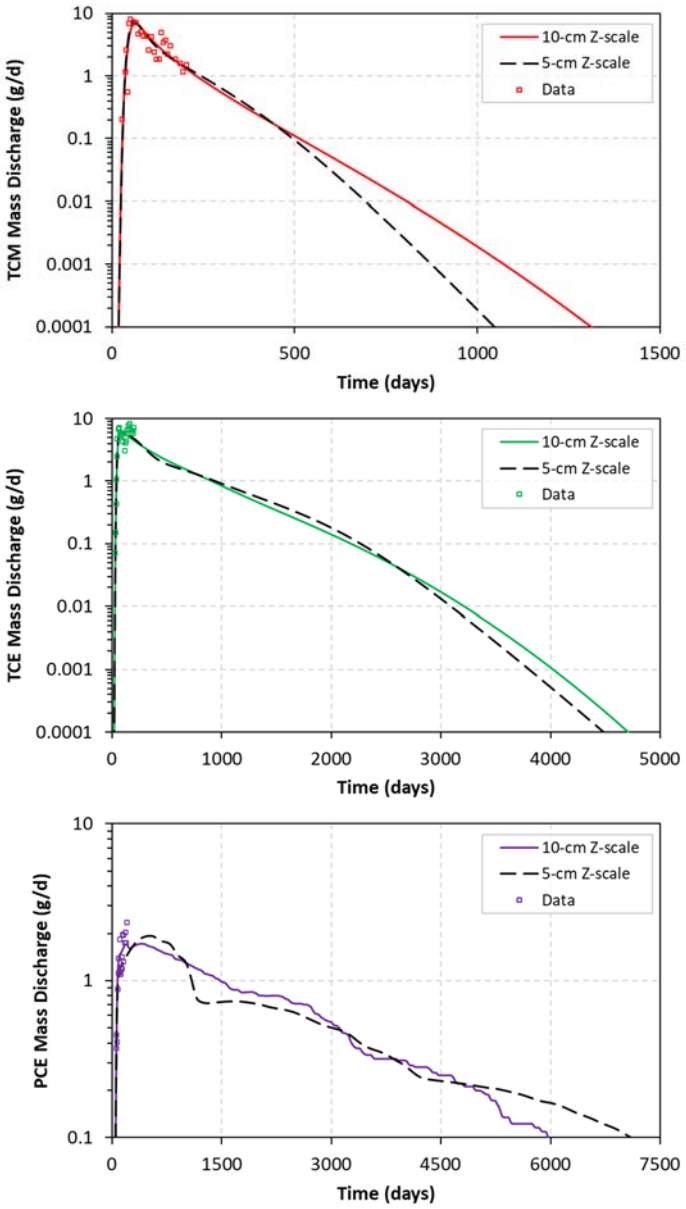

Figure 6. Comparisons of long-term projections of MD profiles generated with the base parameter realization of model C (continuous lines) and the modeling results of Mobile et al. (2012) (dashed lines). All projections were generated with $\beta = 0$. In both cases, multistage behavior of NAPL depletion emerged from the NAPL architecture, which was constrained by the known mass and the post-experiment source footprint in Mobile et al. (2012). Small differences in long-term projections of source depletion emphasized the importance of constraining the source mass.

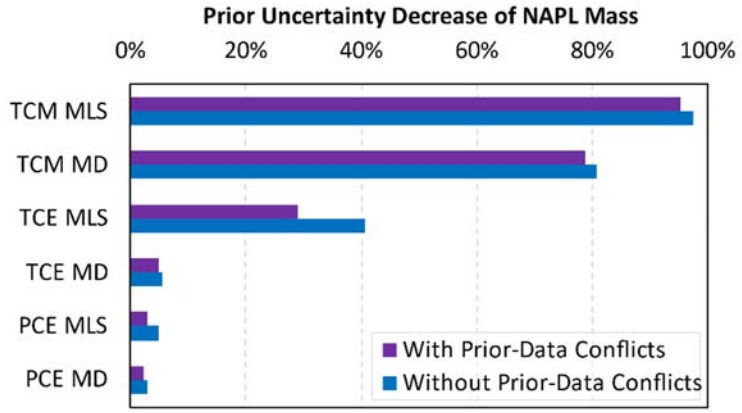

Figure 7. Percent worth of monitoring datasets for reducing the prior uncertainty of initial source mass.

Table 3. Distribution of S_0^N and worth of MLS ports for reducing prior uncertainty of NAPL mass.

NAPL Zone	MLS Port	Prior Uncertainty Variance Decrease	S_0^N (%) A	S_0^N (%) B	S_0^N (%) C
6.1	606	54.0%	9.7%	12.8%	14.0%
6.2	506	42.6%	9.1%	10.4%	11.2%
7.2	507	19.0%	6.2%	6.4%	6.7%
7.1	607	15.7%	6.2%	7.3%	7.5%
5.2	505	5.1%	4.6%	4.6%	5.1%
8.1	608	4.0%	2.6%	2.8%	2.8%
6.3	406	3.2%	0.8%	0.9%	1.0%
7.3	407	2.7%	0.9%	0.9%	0.9%
8.2	508	2.6%	1.2%	1.3%	1.6%
9.3	409	2.4%	2.8%	3.0%	3.1%
9.2	509	1.8%	2.3%	2.4%	2.4%
5.1	605	1.5%	0.1%	0.1%	0.1%
9.1	609	0.7%	0.0%	0.0%	0.0%
9.4	309	0.6%	0.1%	0.0%	0.0%
5.3	405	0.6%	1.0%	1.0%	1.0%
8.3	408	0.3%	0.5%	0.5%	0.7%
10.2	410	0.0%	0.1%	0.1%	0.1%
10.1	510	0.0%	0.2%	0.2%	0.3%
12	412	0.0%	0.1%	0.3%	0.2%
11.1	411	0.0%	0.1%	0.2%	0.2%
11.2	311	0.0%	0.1%	0.1%	0.0%
10.3	310	0.0%	0.1%	0.0%	0.0%
4	404	0.0%	1.1%	1.1%	1.1%

A (PEST_HP, MLS only), B (PESTPP-iES, MLS only), C (PESTPP-iES, MLS and MD)

An example of the worth of individual MLS and MD measurements for reducing the prior uncertainty of NAPL-zone parameters is presented in Figure 8. These results indicated opposite trends in the worth of aqueous-phase concentrations for estimating C_0^N and k_0^N. Maximum

concentrations constrained k_0^N parameters, while declining concentrations constrained C_0^N parameters accounting for NAPL mass. This explained why only TCM data significantly reduced the prior uncertainty of NAPL mass, as all MLS ports showed declining TCM concentrations (Figure 2). Similarly, ~30% (with PDCs) and ~40% (without PDCs) reductions in the prior uncertainty of NAPL mass by TCE MLS data (Figure 7) was attributed to MLS ports with approximately more than 100 g of NAPL mass (e.g., port 508 in Table 2 and Figure 3) and declining TCE concentrations. Conversely, Table 4 indicates parameters with a low RUVR corresponding to deeper NAPL zones accounting for less than 1% of total NAPL mass. The low RUVR values in Table 4 were also caused by narrower prior uncertainty bounds compared to those of other NAPL-zone parameters. As shown in Figure 8, NAPL zones harboring most of the NAPL mass (e.g., 6.1 and 6.2 in Table 3) also benefited from additional C_0^N uncertainty reductions by TCM MD data, highlighting the contribution of those zones to the overall source depletion rates.

Table 4. Source zone parameters with lower than 80% prior uncertainty variance reduction.

MLS	NAPL	RUVR (with PDC)		RUVR (without PDC)	
Port	Zone	C_0^N	$k_{TCM,0}^N$	C_0^N	$k_{TCM,0}^N$
605	5.1	65%	> 80%	79%	> 80%
609	9.1	11%	13%	18%	22%
309	9.4	11%	5%	17%	8%
510	10.1	50%	> 80%	67%	> 80%
410	10.2	> 80%	17%	> 80%	18%
310	10.3	12%	15%	18%	15%
311	11.2	7%	1%	28%	1%
NA	$k_{TCE,0}^N$	NA	5%	NA	8%
NA	$k_{PCE,0}^N$	NA	18%	NA	30%

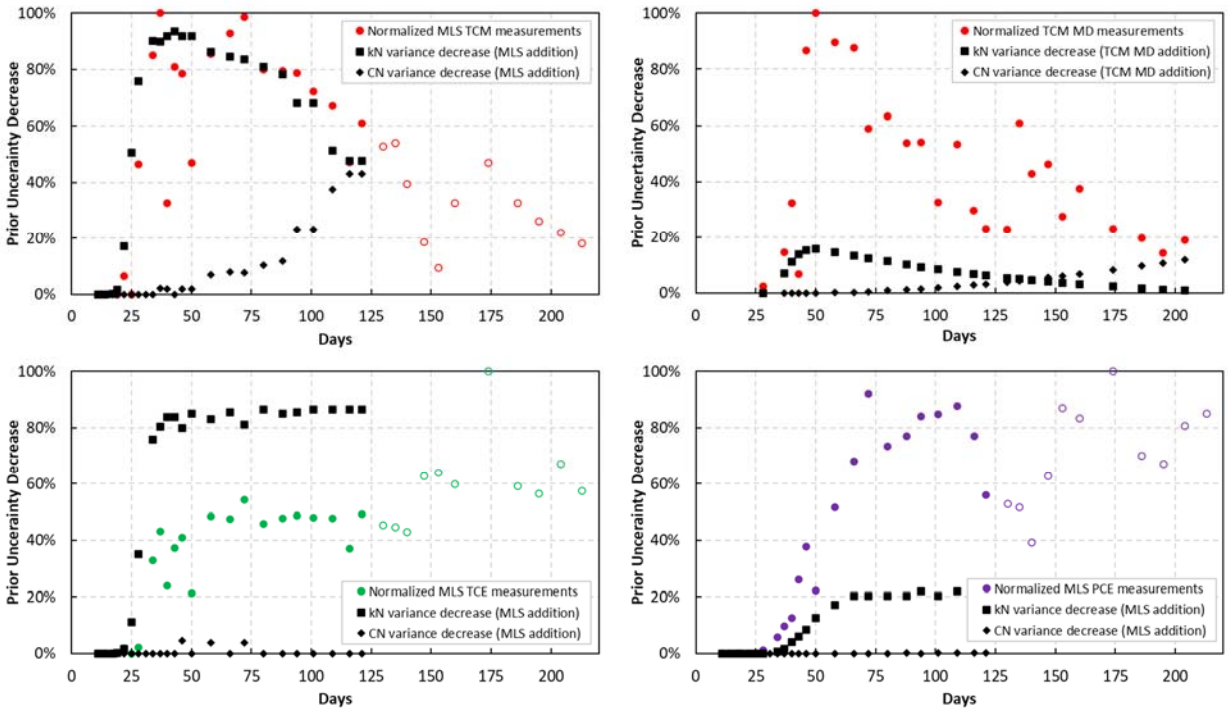

Figure 8. Percent worth of individual aqueous-phase concentrations (MLS port 606 and TCM MD) for reducing the prior uncertainty variance of C_0^N and $k_{TCM,0}^N$ of NAPL zone 6.1. Although the same trend of added value by individual measurements was determined for all ports, only NAPL zones containing most of the source mass benefited from additional uncertainty reductions by TCM MD data. In turn, the correspondence between the TCM MD profile with individual MLS ports emphasized the value of multilevel monitoring for estimating NAPL architecture.

Except for parameters listed in Table 4, history-matching reduced the prior uncertainty of all NAPL parameters by up to 100%. In contrast to C_0^N parameters constrained by TCM data exclusively, maximum TCE and PCE concentrations also constrained $k_{0,TCM}^N$ parameters (Figure 7). However, as indicated in Table 4, the prior uncertainty of $k_{0,TCE}^N$ and $k_{0,PCE}^N$ (global $k_{0,TCM}^N$ multipliers) was not reduced, partially because of their narrow prior uncertainty bounds. Yet the small RUVR of these mass transfer parameters was driven by their corresponding MLS datasets (results not shown). Moreover, while the mean values of both multipliers (Table 1) were in close agreement with those estimated by Mobile et al. (2012), as $k_{TCE,0}^N = 0.96$ and $k_{0,PCE}^N = 0.85$, FOSM analysis with Models A, B, and C suggested that $k_{TCE,0}^N$ could take a value greater than 1, which would not be consistent with previous findings (Mobile et al. 2012) or with mass transfer correlations with component diffusivities (Imhoff et al., 1993; Powers et al., 1992, 1994). As previously discussed, possible explanations for remaining uncertainties on $k_{TCE,0}^N$ and $k_{PCE,0}^N$ could include sub-grid-scale NAPL dissolution behavior, noticeable primarily in TCM MLS data, and/or the influence of grid scale on concentration gradients and $\alpha_{L,T}$ (transverse and longitudinal). Transverse dispersion has been shown to regulate mass transfer rates from DNAPL pools (Hunt & Sitar, 1988, Stewart et al., 2021), requiring an ultrafine grid scale for accurate numerical simulations of DNAPL dissolution (Falta, 2003).

3.3. Analysis of DNAPL Mass Depletion

Table 5 presents mass balance results of NAPL mass removed by natural dissolution calculated using all models. The percent reductions of initial mass were calculated using the known initial

composition values. Despite differences in the estimated source zone architectures, Model C produced nearly identical results as Mobile et al. (2012), emphasizing the importance of constraining NAPL mass for estimating source depletion rates. Conversely, Model A resulted in a 40% reduction of initial NAPL mass, almost doubling Model C results. Model C also indicated a source persistence at the end of the natural dissolution period ~4 and ~2 times lower than indicated by Models A and B, respectively, using PCE as reference. These results reflected the advantage of implementing prior Monte Carlo analyses to understand model deficiencies in relation to the observed system behavior. Additionally, Table 2 indicates the amount of NAPL mass eliminated by methanol remediation, calculated by subtracting the NAPL mass remaining in the soil estimated by Broholm et al. (1999) after conclusion of the experiment, from the remaining mass after 220 days estimated with Model C. Differences in the methanol calculations were linked to post-experiment mass estimated by Broholm et al. (1999) from C_i^N values assuming different S^N values. The obvious impact that such differences would have on remedial designs at hazardous waste sites highlighted difficulties in measuring S^N directly, even by soil confirmatory sampling. These results suggested value in the indirect source characterization method undertaken in this study to estimate and reduce the uncertainty of site-specific mass-transfer parameters, which is critical for effective, risk-based remedial optimizations.

Table 5. Mass of NAPL removed.

Data	NAPL Mass Removed (kg)				Initial NAPL Mass Reduction (%)			
Source	TCM	TCE	PCE	Total	TCM	TCE	PCE	Total
Initial (injected)	0.74	2.92	4.04	7.70	0	0	0	0
Model A	0.67	1.46	0.97	3.10	91	50	24	40
Model B	0.63	1.10	0.55	2.28	85	38	14	30
Model C	0.61	0.92	0.25	1.78	82	32	6	23
Mobile et al. (2012)	0.59	0.91	0.24	1.74	80	31	6	23
Methanol flush (1)	0.13	1.47	1.88	3.49	17	50	47	45
Methanol flush (2)	0.13	1.18	0.80	2.11	17	40	20	27

Mass removed by methanol flushing was calculated by subtracting post-experimental NAPL mass remaining in soil estimated by Broholm et al. (1999) with C_i^N values assuming (1) a homogeneous 3.6% S^N in all excavation layers (Figure S1) and (2) assuming 20% S^N in excavation layer 2 (where a DNAPL pool was observed), from the remaining NAPL mass on day 220 estimated with the posterior base realization of Model C. The percent reductions of initial NAPL mass were calculated with respect to the known initial composition of the mixture.

4. Conclusions

This study demonstrated the worth of high-resolution monitoring and inverse numerical modeling for characterizing a DNAPL source zone. The accuracy of estimated NAPL mass was tied to the depleting signature of MLS and MD aqueous-phase concentrations of the most soluble NAPL component and least by volume (TCM). At contaminated sites, decreasing concentrations may not reflect final NAPL dissolution stages, which could bias estimated parameters and long-term projections of source depletion. The impact of multiscale heterogeneity of NAPL architecture and dissolution on the uncertainty of model parameters was investigated with prior-based Monte Carlo analyses, where PDCs highlighted model limitations for representing sub-grid-scale mass transfer processes. Hence, multiscale heterogeneity of NAPL architecture and dissolution not captured in available monitoring profiles could limit model confidence for remedial-decision making at sites with large and architecturally complex source zones. These situations may benefit from the field

test proposed by Mobile et al. (2016) to determine mass transfer rate coefficients in situ. This test would induce breakthrough of nonequilibrium concentrations through forced hydraulic gradients and flushing in the source zone, generating monitoring profiles suitable for the inverse modeling techniques applied in this study.

Breakthrough data collected from the mass transfer test described in Mobile et al. (2016) would also be useful for allocating grid-scale NAPL zones, similar to the model parameterization guided by MLS data in this study. In contrast to the simplified aquifer parameters in this work, spatially-correlated hydraulic and transport properties can be characterized by geostatistical methods coupled with numerical modeling for management support of source zones in heterogeneous aquifers. Also, high-resolution NAPL delineation with MIP and LIF tools may further reduce uncertainties on source architecture, including residual saturations estimated with inverse numerical modeling. Furthermore, sites where natural attenuation mechanisms are significant may benefit from several monitoring transects along the flow direction to distinguish attenuation capacity from NAPL dissolution rates.

Uncertainty analyses confirmed an inability to estimate the β depletion exponent for any NAPL zone despite declining TCM concentrations measured at all MLS ports. Thus, predictive timeframes of source mass depletion should include a variability range for β parameters with ensemble realizations, representing the transient nature of NAPL mass transfer rates in a stochastic manner. Several advantages of parameter optimization and uncertainty quantification with model ensembles were also demonstrated in this investigation. For example, while single parameter sets per each NAPL zone could not represent sub-grid-scale multistage dissolution profiles, removing early-stage TCM peak concentrations reduced the propensity for biasing mass transfer rates and improved the accuracy of NAPL mass estimation. Accuracy of mass estimation was also attained through parameter parsimony, as estimating NAPL-zone parameters without directly upgradient MLS ports produced inconsistent and erroneous results. The importance of accurately constraining the source mass was emphasized in a similar projection of source depletion compared to Mobile et al. (2012), where multistage NAPL dissolution behavior in both models emerged from MLS constraints.

Prior-based Monte Carlo and FOSM analyses suggested that simulating interphase mass transfer from NAPL mixtures may be influenced by grid scale, despite incorporating adjustable compound-specific mass transfer rate coefficients. Specifically, FOSM results indicated no prior uncertainty reductions on the global multipliers of mass transfer rate coefficients, while $k_{TCE,0}{}^{N}$ showed a tendency to exceed its upper uncertainty limit in all models, potentially biasing source dissipation timeframes. Although these results did not prevent a reasonable estimation of initial source mass and consistent ranges of mass transfer rates, further research is required to investigate the impact of grid scale on dispersivity and mass transfer rate coefficients describing multicomponent NAPL dissolution of source zones with heterogeneous architectures (i.e., comprised by ganglia- and pool-dominated accumulations of NAPL mass). As demonstrated, combining mass discharge/flux rates with high-resolution monitoring can improve history-matching of noisy data, where ensemble-based parameter estimation considering measurement noise can reduce parameter bias without resorting to a more complex simulation of multiple subsurface processes, supporting the indirect characterization of NAPL source zones.

Acknowledgements

This study was partially supported by the Environmental Security Technology Certification Program (ESTCP) under Project ER19-5223. The content of this manuscript has not been subject to agency review and does not necessarily represent the view of the sponsoring agency.

Data Availability Statement

The Groundwater Modeling System (GMS) software hosting the MODFLOW2000 and SEAM3D programs used in this study is available through Aquaveo at https://www.aquaveo.com/software/gms-groundwater-modeling-system-introduction. The PWTADJ2, PREDUNC, GENLINPRED, and PEST_HP programs are available at https://pesthomepage.org/programs. The PESTPP-iES software is available at https://www.usgs.gov/software/pest-software-suite-parameter-estimation-uncertainty-analysis-management-optimization-and (version 5.1.6 was used and the source code is available on https://github.com/usgs/pestpp/releases/tag/5.1.6). Figures 2 through 4 were produced with the Matplotlib (https://matplotlib.org/) version 3.5.1 and Seaborn (https://seaborn.pydata.org/) version 0.11.2 libraries using the Python programming language. Aqueous-phase concentration data will be archived in an online repository maintained by Virginia Tech with a unique DOI number.

5. References

Abriola, L. M. (1989). Modeling Multiphase Migration of Organic Chemicals in Groundwater Systems - A Review and Assessment. *Environmental Health Perspectives, 8*, 117-143. https://doi.org/10.1289/ehp.8983117

Abriola, L. M., Miller, E. L., Pennell, K. D., Ramsburg, A., & Christ, J. A. (2013). *Metric identification and protocol development for characterizing DNAPL source zone architecture and associated plume response.* Alexandria, VA: SERDP Project ER-1612.

Agaoglu, B., Copty, N. K., Scheytt, T., & Hinkelmann, R. (2015). Interphase mass transfer between fluids in subsurface formations: A review. *Advances in Water Resources, 79*, 162-194. https://doi.org/10.1016/j.advwatres.2015.02.009

Arshadi, M., De Paolis Kaluza, M. C., Miller, E. L., & Abriola, L. M. (2020). Subsurface Source Zone Characterization and Uncertainty Quantification Using Discriminative Random Fields. *Water Resources Research*. https://doi.org/10.1029/2019WR026481

Broholm, K. (2007). Solvent release into a sandy aquifer 3: enhanced dissolution by methanol injection. *Environmental Technology, 28*(1), 11-18. https://doi.org/10.1080/09593332808618767

Broholm, K., Feenstra, S., & Cherry, J. A. (1999). Solvent Release into a Sandy Aquifer. 1. Overview of Source Distribution and Dissolution Behavior. *Environmental Sicence & Technology, 33*(5), 681-690. https://doi.org/10.1021/es980097d

679 Broholm, K., Feenstra, S., & Cherry, J. A. (2005). Solvent Release into a Sandy Aquifer. 2.
680 Estimation of DNAPL Mass Based on a Multiple-Component Dissolution Model.
681 *Environmental Science & Technology, 39*(1), 317-324. https://doi.org/10.1021/es0306462

682 Brusseau, M. L., Matthieu III, D. E., Carroll, K. C., Mainhagu, J., Morrison, C., McMillan, A., . .
683 . Plaschke, M. (2013). Characterizing Long-term Contaminant Mass Discharge and the
684 Relationship Between Reductions in Discharge and Reductions in Mass for DNAPL
685 Source Areas. *Journal of Contaminant Hydrology*, 1-12.
686 https://doi.org/10.1016/j.jconhyd.2013.02.011

687 Christ, J. A., Ramsburg, A. C., Pennell, K. D., & Abriola, L. M. (2006). Estimating mass discharge
688 from dense nonaqueous phase liquid source zones using upscaled mass transfer
689 coefficients: An evaluation using multiphase numerical simulations. *Water Resources*
690 *Research, 42*(11). https://doi.org/10.1029/2006WR004886

691 Christ, J. A., Ramsburg, C. A., Pennell, K. D., & Abriola, L. M. (2010). Predicting DNAPL mass
692 discharge from pool-dominated source zones. *Journal of Contaminant Hydrology, 114*(1-
693 4), 18 - 34. https://doi.org/10.1016/j.jconhyd.2010.02.005

694 Dekker, T. J., & Abriola, L. M. (2000). The influence of field-scale heterogeneity on the infiltration
695 and entrapment of dense nonaqueous phase liquids in saturated formations. *Journal of*
696 *Contaminant Hydrology, 42*(2-4), 187-218. https://doi.org/10.1016/S0169-
697 7722(99)00092-3

698 DiFilippo, E. L., & Brusseau, M. L. (2008). Relationship Between Mass Flux Reduction and
699 Source-Zone Mass Removal: Analysis of Field Data. *Journal of Contaminant Hydrology,*
700 *98*(1-2), 22-35. https://doi.org/10.1016/j.jconhyd.2008.02.004

701 DiFilippo, E. L., & Brusseau, M. L. (2011). Assessment of a Simple Function to Evaluate the
702 Relationship Between Mass Flux Reduction and Mass Removal for Organic-Liquid
703 Contaminated Source Zones. *Journal of Contaminant Hydrology, 123*(3-4), 104-113.
704 https://doi.org/10.1016/j.jconhyd.2010.12.011

705 Doherty, J. (2015). *Calibration and Uncertainty Analysis for Complex Environmental Models.*
706 Brisbane, Australia: Watermark Numerical Computing.

707 Doherty, J. (2020). *PEST_HP. PEST for highly parallelized computing environments.* Watermark
708 Numerical Computing. Retrieved from https://pesthomepage.org/documentation

709 Engelmann, C., Handel, F., Binder, M., Yadav, P. K., Dietrich, P., Liedl, R., & Walther, M. (2019).
710 The fate of DNAPL contaminants in non-consolidated subsurface systems – Discussion on
711 the relevance of effective source zone geometries for plume propagation. *Journal of*
712 *Hazardous Materials, 375*, 233-240.
713 https://doi.org/doi.org/10.1016/j.jhazmat.2019.04.083

714 Falta, R. (2003). Modeling sub-grid-block-scale dense nonaqueous phase liquid (DNAPL) pool
715 dissolution using a dual-domain approach. *Water Resources Research, 39*(12).
716 https://doi.org/10.1029/2003WR002351

Falta, R. W., Rao, S. P., & Basu, N. (2005a). Assessing the impacts of partial mass depletion in DNAPL source zones I. Analytical modeling of source strength functions and plume response. *Journal of Contaminant Hydrology*, 259-280. https://doi.org/0.1016/j.jconhyd.2005.05.010

Falta, R. W., Basu, N., & Rao, S. P. (2005b). Assessing impacts of partial mass depletion in DNAPL source zones: II. Coupling source strength functions to plume evolution. *Journal of Contaminant Hydrology*, 45-66. https://doi.org/10.1016/j.jconhyd.2005.05.012

Frind, E. O., Molson, J. W., & Schirmer, M. (1999). Dissolution and mass transfer of multiple organics under field conditions: The Borden emplaced source. *Water Resources Research, 35*(3), 683-694. https://doi.org/10.1029/1998WR900064

Geller, J. T., & Hunt, J. R. (1993). Mass Transfer From Nonaqueous Phase organic Liquids in Water-Saturated Porous Media. *Water Resources Research, 29*(4), 883-845. https://doi.org/10.1029/92WR02581

Guo, Z., Russo, A. E., DiFilippo, E. L., Zhang, Z., Zheng, C., & Brusseau, M. L. (2020). Mathematical modeling of organic liquid dissolution in heterogeneous source zones. *Journal of Contaminant Hydrology, 235*. https://doi.org/10.1016/j.jconhyd.2020.103716

Harbaugh, A. W., Banta, E. R., Hill, M. C., & McDonald, M. G. (2000). *MODFLOW-2000, The U.S. Geological Survey Modular Ground-Water Model: User Guide to Modularization Concepts and the Ground-Water Flow Process.* Reston, Virginia: U.S. Geological Survey. https://doi.org/10.3133/ofr200092

Hunt, J. R., & Sitar, N. (1988). Nonaqueous Phase Liquid Transpot and Cleanup 1. Analysis of Mechanisms. *Water Resources Research, 24*(8), 1247-1258. https://doi.org/10.1029/WR024i008p01247

Imhoff, P. T., Jaffe, P. R., & Pinder, G. F. (1993). An experimental study of complete dissolution of a nonaqueous phase liquid in saturated porous media. *Water Resources Research, 30*(2), 307-320. https://doi.org/10.1029/93WR02675

Kang, X., Kokkinaki, A., Kitandis, P. K., Shi, X., Lee, J., Mo, S., & Wu, J. (2021a). Hydrogeophysical Characterization of Nonstationary DNAPL Source Zones by Integrating a Convolutional Variational Autoencoder and Ensemble Smoother. *Water Resources Research, 57*(1). https://doi.org/10.1029/2020WR028538

Kang, X., Kokkinaki, A., Power, C., Kitandis, P. K., Shi, X., Duan, L., . . . Wu, J. (2021b). Integrating deep learning-based data assimilation and hydrogeophysical data for improved monitoring of DNAPL source zones during remediation. *Journal of Hydrology, 601*, 126655. https://doi.org/10.1016/j.jhydrol.2021.126655

Koch, J., & Nowak, W. (2015). Predicting DNAPL mass discharge and contaminated site longevity probabilities: Conceptual model and high-resolution stochastic simulation. *Water Resources Research, 51*(2), 806 - 831. https://doi.org/10.1002/2014WR015478.

Koch, J., & Nowak, W. (2016). Identification of contaminant source architectures—A statistical inversion that emulates multiphase physics in a computationally practicable manner. *Water Resources Research, 52*, 1009–1025. https://doi.org/10.1002/2015WR017894

Kokkinaki, A., O'Carroll, M., Werth, C. J., & Sleep, B. E. (2013). Coupled simulation of DNAPL infiltration and dissolution in three-dimensional heterogeneous domains: Process model validation. *Water Resources Research, 49*, 7023-7036. https://doi.org/10.1002/wrcr.20503, 2013

Kokkinaki, A., Werth, C. J., & Sleep, B. E. (2014). Comparison of upscaled models for multistage mass discharge from DNAPL source zones. *Water Resources Research*, 3187 - 3205. https://doi.org/10.1002/2013WR014663

Kueper, B. H., Stroo, H. F., Vogel, C. M., & Ward, C. H. (2014). *Chlorinated Solvent Source Zone Remediation.* Springer New York. https://doi.org/10.1007/978-1-4614-6922-3

Luciano, A., Mancini, G., Torreta, V., & Viotti, P. (2018). An empirical model for the evaluation of the dissolution rate from a DNAPL-contaminated area. *Environmental Science and Pollution Research*, 33992-34004. https://doi.org/10.1007/s11356-018-3193-6

Marble, J. C., DiFilippo, E. L., Zhang, Z., Tick, G. R., & Brusseau, M. L. (2008). Application of a lumped-process mathematical model to dissolution of non-uniformly distributed immiscible liquid in heterogeneous porous media. *Journal of Contaminant Hydrology, 100*, 1-10. https://doi.org/10.1016/j.jconhyd.2008.04.003

Mayer, A. S., & Hassanizadeh, M. S. (2005). *Soil and Groundwater Contamination: Nonaqueous Phase Liquids–Principles and Observations.* Washington, D.C.: American Geophysical Union.

McMillan, L. A., Rivett, M. O., Wealthall, G. P., Zeeb, P., & Dumble, P. (2018). Monitoring well utility in a heterogeneous DNAPL source zone area: Insights from proximal multilevel sampler wells and sampling capture-zone modelling. *Journal of Contaminant Hydrology, 210*, 15-30. https://doi.org/10.1016/j.jconhyd.2018.02.001

Miller, C. T., Christakos, G., Imhoff, P. T., McBride, J. F., & Pedit, J. A. (1998). Multiphase flow and transport modeling in heterogeneous porous media: challenges and approaches. *Advances in Water Resources, 21*(2), 77-120. https://doi.org/10.1016/S0309-1708(96)00036-X

Mobile, M. A., Widdowson, M. A., & Gallagher, D. L. (2012). Multicomponent NAPL Source Dissolution: Evaluation of Mass-Transfer Coefficients. *Environmental Science & Technology, 46*(18), 10047-10054. https://doi.org/10.1021/es301076p

Mobile, M., Widdowson, M., Stewart, L., Nyman, J., Deeb, R., Kavanaugh, M., . . . Gallagher, D. (2016). In-situ determination of field-scale NAPL mass transfer coefficients: Performance, simulation and analysis. *Journal of Contaminant Hydrology, 187*, 31-46. https://doi.org/10.1016/j.jconhyd.2016.01.010

National Research Council. (2005). *Contaminants in the subsurface: Source zone assessment and remediation.* Washington, D.C.: The National Academic Press.

Park, E., & Parker, J. C. (2005). Evaluation of an upscaled model for DNAPL dissolution kinetics in heterogeneous aquifers. *Advances in Water Resources*, 1280-1291. https://doi.org/10.1016/j.advwatres.2005.04.002

Powers, S. E., Abriola, L. M., & Weber Jr, W. J. (1992). An Experimental Investigation of Nonaqueous Phase Liquid Dissolution in Saturated Subsurface Systems: Steady State Mass Transfer Rates. *Water Resources Research, 28*(10), 2691-2705. https://doi.org/10.1029/92WR00984

Powers, S. E., Abriola, L. M., & Weber, W. J. (1994). An experimental investigation of nonaqueous phase liquid dissolution in saturated systems: Transient mass transfer rates. *Water Resources Research, 30*(2), 321-332. https://doi.org/10.1029/93WR02923

Rivett, M. O., & Feenstra, S. (2005). Dissolution of an Emplaced Source of DNAPL in a Natural Aquifer Setting. *Environmental Science & Technology, 39*, 447-455. https://doi.org/10.1021/es040016f

Saenton, S., & Illangasekare, T. H. (2004). Determination of DNAPL entrapment architecture using experimentally validated numerical codes and inverse modeling. *Developments in Water Science, 55*, 767-778. https://doi.org/10.1016/S0167-5648(04)80098-4

Saenton, S., & Illangasekare, T. H. (2007). Upscaling of mass transfer rate coefficient for the numerical simulation of dense nonaqueous phase liquid dissolution in heterogeneous aquifers. *Water Resources Research, 43*(2). https://doi.org/10.1029/2005WR004274

Soerens, T. S., Sabatini, D. A., & Harwell, J. H. (1998). Effects of flow bypassing and nonuniform NAPL distribution on the mass transfer characteristics of NAPL dissolution. *Water Resources Research, 34*(7), 1657-1673. https://doi.org/10.1029/98WR00554

Stewart, L. D., Chambon, J. C., Widdowson, M. A., & Kavanaugh, M. C. (2022). Upscaled modeling of complex DNAPL dissolution. *Journal of Contaminant Hydrology, 244*. https://doi.org/10.1016/j.jconhyd.2021.103920

Waddill, D. W., & Widdowson, M. A. (2000). *SEAM3D: A numerical model for three-dimensional solute transport and sequential electron acceptor-based bioremediation in groundwater.* ERDC/EL TR- 00-18. U.S. Army Engineer Research and Development Center, Vicksburg, MS.

Watermark Numerical Computing. (2018). *Model-Independent Parameter Estimation. User Manual Part II: PEST Utility Suport Software.* Retrieved from https://pesthomepage.org/documentation

White, J. T. (2018). A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions. *Environmental*

Modelling & Software, 109, 191-201. https://doi.org/doi.org/10.1016/j.envsoft.2018.06.009

White, J. T., Hemmings, B., Fienen, M. N., & Knowling, M. J. (2021). Towards improved environmental modeling outcomes: Enabling low-cost access to high-dimensional, geostatistical-based decision-support analyses. *Environmental Modelling and Software, 139*, 105022. https://doi.org/10.1016/j.envsoft.2021.105022

White, J., Hunt, R., Fienen, M., & Doherty, J. (2020). *Approaches to Highly Parameterized Inversion: PEST++ Version 5, a Software Suite for Parameter Estimation, Uncertainty Analysis, Management Optimization and Sensitivity Analysis.* Reston, VA: U.S. Geological Survey. https://doi.org/10.3133/tm7C26

Yang, L., Wang, X., Mendoza-Sanchez, I., & Abriola, L. M. (2018). Modeling the influence of coupled mass transfer processes on mass flux downgradient of heterogeneous DNAPL source zones. *Journal of Contaminant Hydrology, 211*, 1-14. https://doi.org/10.1016/j.jconhyd.2018.02.003