Quantifying DNAPL source zone longevity with upscaled modeling: practical insights from flow-cell experiments and uncertainty analyses

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

- Upscaled modeling and uncertainty analyses of flow-cell experiments elucidated upon data assimilation strategies at DNAPL sites
- Parameterization of source zone heterogeneities for history-matching was necessary to predict unbiased DNAPL dissolution timeframes
- Coarse DNAPL delineation sufficed to quantify unbiased uncertainty limits of source zone lifespans a priori

Abstract

Estimating dissipation timeframes and contaminant mass discharge rates of dense non-aqueous phase liquids (DNAPLs) source zones is of key interest for environmental-management support. Upscaled mathematical modeling of DNAPL dissolution provides a practical approach for assimilating site characterization and downgradient monitoring data to constrain future system behavior. Yet significant uncertainties on predictions of source zone dissipation rates may arise from inadequate or inaccurate conceptual assumptions in parameterization designs. These implications were investigated through upscaled modeling, sensitivity, and uncertainty analyses of high-resolution flow-cell experiments. Sensitivity results emphasized the role of local groundwater velocity and source dimensions in mass transfer scaling by strongly influencing error with respect to DNAPL persistence and dissolution rates. Linear uncertainty analyses, facilitated by PEST ancillary software, demonstrated the worth of monitoring profiles for constraining DNAPL saturations and dispersive mass transfer rates, responsible for source zone longevity. Nonlinear analyses performed with the iterative ensemble smoother PESTPP-iES, facilitated the quantification of unbiased source dissipation uncertainties from DNAPL delineation data. Conversely, monitoring data assimilation without consideration of flow-field heterogeneity and saturation distribution along the flow path biased model predictions. Our analyses provided practical recommendations on upscaled model design to assimilate available site data and support remedial-decision making.

Plain Language Summary

Currently, remedial-decision makers rarely benefit from physically-based modeling methods and uncertainty analyses to manage sites impacted by DNAPL source zones. Difficulties in estimating

DNAPL dissolution rates stem from source zone heterogeneities, which are difficult to characterize in detail, compounded by a lack of scalable methodologies connecting source zone characterization with discharge monitoring. In addition, monitoring and site characterization efforts supporting performance-based remedial objectives are typically uninformed by uncertainty evaluations predicated on DNAPL mass transfer processes. To bridge that gap, we investigated the impact of data-driven conceptual assumptions on predictions of source zone behavior by coupling a practical DNAPL dissolution model with uncertainty quantification methods. Simulations of flow-cell experiments demonstrated the worth of DNAPL delineation for constraining source zone dissipation uncertainties, estimated a priori through parameterization of DNAPL distributions. In turn, parameterizing system heterogeneities in greater detail was necessary to estimate unbiased source zone characteristics and lifespans, derived from the assimilation of complex DNAPL dissolution trends observed in monitoring profiles. Our results demonstrated how available site data can be integrated into a decision-support modeling framework to inform data collection strategies and remedial designs.

Index Terms and Keywords

Source zone persistence, source zone heterogeneity, DNAPL dissolution rates, conceptual assumptions, data assimilation, model parameterization, remedial-decision making, practical recommendations

1. Introduction

Mathematical modeling can provide valuable insights for decision support at hazardous waste sites with groundwater impacted by dense non-aqueous phase liquids (DNAPLs). However, a gap between simplistic analytical screening models and overly complex numerical simulators has limited their applicability for estimating DNAPL longevity and dissolution rates. Researchers have focused on estimating distributions of DNAPL saturation, referred to as the source zone architecture, or DNAPL dissolution rates from synthetically-generated datasets 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; Powers et al., 1992, 1994; Saenton & Illangasekare, 2004).

Decision-support modeling incorporating LEA is questionable because heterogeneity of aquifer hydraulic properties and source 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). 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 properties (Powers et al., 1992; Saenton & Illangasekare, 2007). Moreover, the computational cost of pore-scale numerical models incorporating LEA and Gilland-Sherwood correlations limits their practicality for data assimilation and uncertainty quantification (Falta, 2003; Kokkinaki et al., 2013; Powers et al., 1994). An alternative method is predicated on a lumped-process, scale-dependent mass transfer coefficient estimated from monitoring data (Guo et al., 2020; Mobile et al., 2012; Park & Parker, 2005). However, estimating mass transfer rates exclusively from

historical monitoring may bias predictions of source longevity because of architectural changes. For example, early in the life cycle of a DNAPL source zone, the contributions of slowly dissolving pools governing complete depletion may not be discernible in discharge data (Abriola et al., 2013).

Multistage DNAPL dissolution, typically observed at contaminated sites and in dissolution experiments, arises from heterogeneity of source zone architecture (Figure 1) primarily comprised by residual ganglia and high-saturation DNAPL pools (Christ et al., 2010; Dekker & Abriola, 2000; DiFilippo & Brusseau, 2008; Lemke & Abriola, 2006; Parker & Park, 2004). Consequently, a number of high-resolution site characterization (HRSC) technologies have been developed (Einarson et al., 2018; Horst et al., 2018; Kueper et al., 2014). Delineation of DNAPLs comprised by chlorinated ethenes is possible with dye-enhanced laser induced fluorescence (DyeLIF) and confirmatory sampling, and with indirect observation methods, such as multilevel monitoring and groundwater extraction systems, passive flux meters, push-pull tracer tests, etc. (Horst et al. 2018; Huang et al. 2010; ITRC, 2010; Kueper et al., 2014). Although HRSC may help constrain DNAPL distribution, quantifying residual mass and saturation directly is not possible (Einarson et al., 2018). Thus, inverse modeling techniques have been applied to quantify lumped-process mass transfer coefficients from monitoring data Marble et al., 2008; Mobile et al., 2012; Saenton & Illangasekare, 2004), or to estimate source zone architectures from borehole and geophysical measurements using physically-based parameterization mechanisms (Kang et al. 2021a, 2021b). Kang et al. (2022) demonstrated a novel Bayesian inversion framework to reconstruct complex permeability and DNAPL saturation fields, subsequently parameterizing an upscaled model of DNAPL dissolution to reproduce experimental source depletion trends.

Upscaled (domain-averaged) models lacking a physical mass transfer basis cannot assimilate HRSC data and have proved ineffective at explaining and predicting DNAPL dissolution behavior (Christ et al., 2006; Kokkinaki et al., 2014; Marble et al., 2008). However, upscaled models incorporating metrics describing the source zone architecture, such as the ganglia-to-pool (GTP) mass ratio, have shown improved success (Abriola et al. 2013; Christ et al., 2010; DiFilippo & Brusseau, 2011). Stewart et al. (2021) developed a volume-averaged (VA) model of DNAPL dissolution predicated on characteristic length scales of DNAPL accumulations and their relative location along groundwater flow paths, explicitly accounting for mass transfer processes at the source zone scale (Figure 1). The model accurately reproduced complex DNAPL dissolution observed in laboratory, numerical, and field experiments by parameterizing initial, and estimable system characteristics without undertaking history-matching. The VA model is therefore able to assimilate HRSC and/or monitoring data to estimate source dissipation timeframes with computational efficiency in a scalable and physically-based manner. Such capabilities make the VA model suitable for evaluating site conceptual assumptions and quantifying uncertainties, which is necessary for effective remedial-decision support (Abriola, 2005).

The primary objectives in this work were to (i) identify the relative contribution of VA mass transfer parameters to source zone dissipation uncertainties, and (ii) investigate how model parameterization influences predictive bias through monitoring data assimilation. The VA DNAPL dissolution model developed by Stewart et al. (2021) was coupled with sensitivity and uncertainty analysis methods to evaluate the worth of direct and indirect source zone measurements for constraining system parameters and model uncertainty. High-resolution datasets of two DNAPL dissolution experiments were leveraged to evaluate data-driven conceptual assumptions on

modeling outcomes. Our findings elucidate on model design to quantify unbiased DNAPL persistence uncertainties, yielding recommendations on HRSC and monitoring data assimilation for constraining future source zone behavior.

2. Upscaled and Volume-Averaged Model of DNAPL Dissolution

Volume-averaging relaxes the need to specify precise locations of DNAPL accumulations within a finely discretized domain. The approach facilitates the incorporation of physically-based mass transfer relationships for complex field-scale dissolution behavior with computational efficiency. As presented by Stewart et al. (2021), dissolution of a single DNAPL accumulation, defined as a volume of relative uniformity in saturation, can be simulated by a generalized upscaled mass transfer function:

$$k_a = \frac{U_0}{V_S}\left[A_{a,yz}k_r(m_a) + A_{a,xy}\sqrt{\frac{4\alpha_T}{\pi X_a}}\right]\left(\frac{m_a}{m_{a,0}}\right)^{\gamma} \quad (1)$$

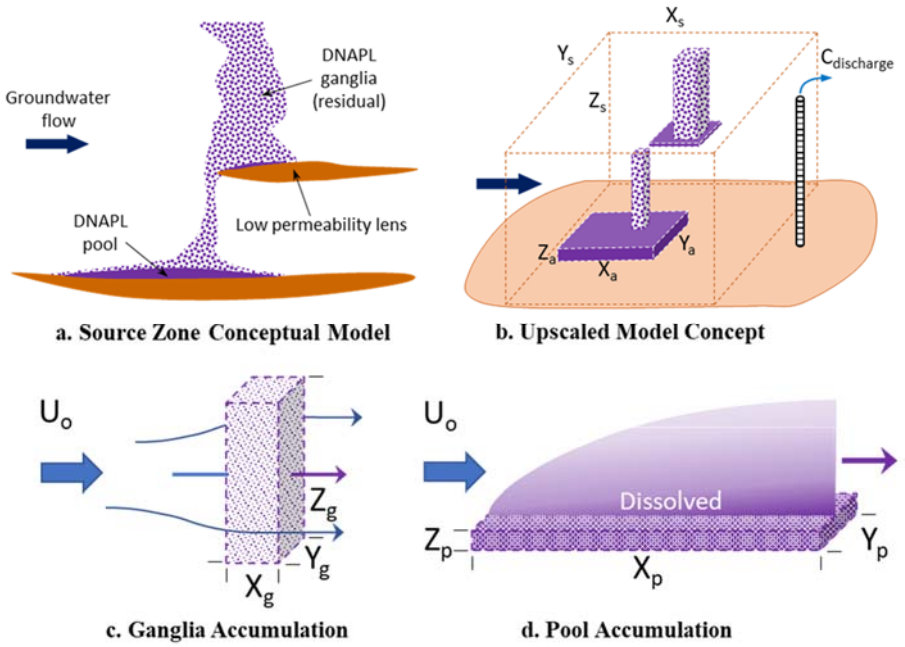

Figure 1. Conceptual and volume-averaged model representations (a and b, respectively) of a DNAPL source zone comprised by characteristic accumulations of (c) residual ganglia and (d) pools. Adapted from Stewart et al. (2021).

Where interphase mass transfer (k_a^N) from an individual DNAPL accumulation *"a"* is driven by the local Darcy groundwater velocity (U_0) upscaled by a source zone reference volume (V_S) encompassing the DNAPL masses. The term on the left inside the brackets represents dissolution attributable to through flow (Figure 1c), which is proportional to the projected area facing flow ($A_{a,yz} = Y_aZ_a$) of "*a*". Flow through "*a*" is regulated by the soil relative permeability (k_r) which gradually increases the DNAPL dissolution rate as the DNAPL volume is reduced. The term on the right represents dissolution attributable to dispersion into bypassing flow (Figure 1d), which is proportional to the hydrodynamic transverse dispersivity (α_T) around "*a*" and the horizontal area of the accumulation ($A_{a,xy} = X_aY_a$). Mass dissolution from low DNAPL saturations, i.e., ganglia,

are dominated by flow through, while high saturation zones, i.e., pools, can be dominated by dispersion. The normalized mass (m/m_0) term reflects a transient reduction of DNAPL interfacial area. Theoretically, the dimensionless exponent $\gamma = 2/3$ for ganglia-dominated accumulations, and $\gamma = ½$ for pool-dominated accumulations. The γ exponent may be adjusted during model history-matching to allow for deviations from conceptual mass transfer assumptions but is expected to fall within this relatively narrow range (Stewart et al., 2021).

2.1. Simulation of DNAPL Dissolution Experiments

Flow-cell experiments presented by DiFilippo et al. (2010) and analyzed by Guo et al. (2020) with a simplified inverse modeling method were utilized in this study. Stewart et al. (2021) simulated these experiments with the VA approach by explicitly accounting for DNAPL saturation distributions, flow field characteristics, and soil properties. The dissolution experiments consisted of two source zone scenarios: a "mixed" DNAPL architecture comprised by a ganglia-dominated accumulation and a pool-dominated accumulation in homogeneous sand, and multiple DNAPL accumulations in a "heterogeneous" soil. Details of model conceptualization and simulation results were presented in Stewart et al. (2021).

2.1.1. Mixed DNAPL Architecture

The "mixed" source zone experiment conducted by DiFilippo et al. (2010) consisted of a uniform pack of sand (40/50 mesh) with a 2-cm thick capillary barrier located along the bottom of the test cell (Figure 2a). An injection of ~12 milliliters of trichloroethene (TCE) at the top of the test cell followed by 48-hour period prior to flow initiation generated a stable source zone architecture consisting of a vertical ganglia zone underlain by a pool. The DNAPL saturation distribution was characterized using a light reflection visualization (LRV) method and TCE effluent concentrations were monitored until source zone depletion.

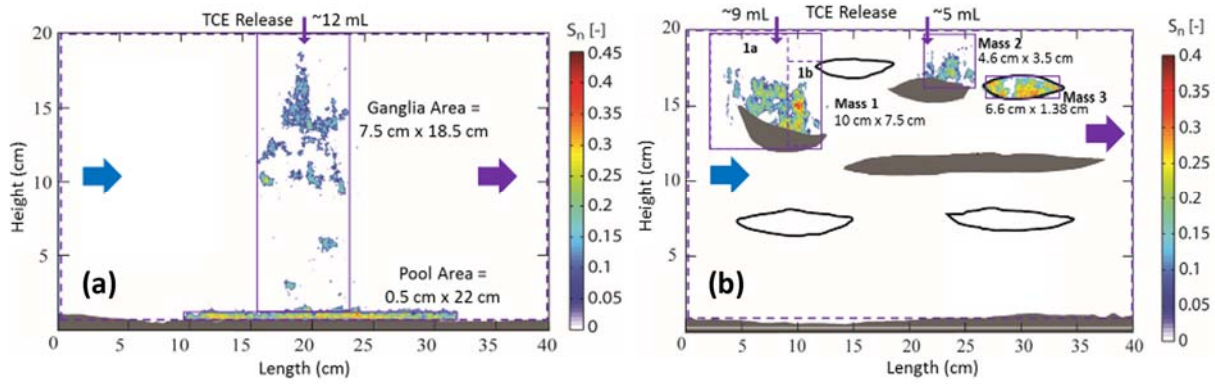

Figure 2. Model conceptualizations of the flow-cell experiments: (a) mixed source zone architecture and (b) heterogeneous source zone. Sub-volumes (purple rectangles) correspond to DNAPL accumulations with distinct saturations encompassed by the source volume (purple dashed line). Adapted from Stewart et al. (2021).

2.1.2. Multiple DNAPL accumulations in Heterogeneous Soil

The test cell of the heterogeneous source experiment (Figure 2b) consisted of a matrix of homogeneous sand (40/50 mesh) with coarser (20/30 mesh) and finer (70/100 mesh) lenticular

zones (DeFilippo et al., 2010; Guo et al., 2020). An injection of ~15 milliliters of TCE at the top of the cell was distributed between two ports with 66% in the far left (upgradient) port and 33% in the center (downgradient) port (DeFilippo et al. 2010). The central release generated two distinct accumulations: one above a fine-grained lens and one within a coarse-grained lens. The coarse lens had an intrinsic permeability approximately 3.5 times higher than the surrounding bulk sands (DeFilippo et al., 2010), resulting in a higher velocity through this material than in the surrounding matrix. As shown in Figure 2b, Stewart et al. (2021) subdivided the upgradient accumulation into two accumulations on the basis of characteristic saturations to accurately capture the measured TCE effluent breakthrough. Sequential dissolution inhibition was also implemented by Stewart et al. (2021) for the two downgradient accumulations on the basis of their relative locations along the flow path. Nomenclature for variables in the model are presented in Table 1.

Table 1. Nomenclature of input parameters used in the VA model of the flow-cell experiments.

Parameter	**Mixed Source**		**Heterogeneous Source**				**Unit**
Mass	*Mass G.*	*Mass P.*	*Mass 1A*	*Mass 1B*	*Mass 2*	*Mass 3*	g
Length	X_g	X_p	X_{1A}	X_{1B}	X_2	X_3	m
Width	Y_g	Y_p	Y_{1A}	Y_{1B}	Y_2	Y_3	m
Height	Z_g	Z_p	Z_{1A}	Z_{1B}	Z_2	Z_3	m
NAPL Saturation	S_g^N	S_p^N	S_{1A}^N	S_{1B}^N	S_2^N	S_3^N	%
Area Facing Flow	YZ_g	YZ_p	YZ_{1A}	YZ_{1B}	YZ_2	YZ_3	m^2
Dispersive Area	XY_g	XY_p	XY_{1A}	XY_{1B}	XY_2	XY_3	m^2
Dispersivity	$\alpha_{T,g}$	$\alpha_{T,p}$	$\alpha_{T,1A}$	$\alpha_{T,1B}$	$\alpha_{T,2}$	$\alpha_{T,3}$	m
γ	γ^g	γ^p	γ^{1A}	γ^{1B}	γ^2	γ^3	-

2.2. Sensitivity Analysis

Model output variability was evaluated with local sensitivity analysis by systematically perturbing input parameters around reference values conceptualized in Stewart et al. (2021). The goal was to compare relative sensitivities with respect to measured discharge concentrations and with respect to the time required to reach cleanup concentrations, defined here as time of remediation (TOR). Both metrics were evaluated using the same model input variability around base parameter sets. Because the plausible variability range of some parameters and their corresponding outputs differs by orders of magnitude compared to those of other parameters, sensitivity coefficients were scaled by maximum values to provide a relative comparison metric of simulation error. All sensitivity analyses were automated coupling SENSAN and PEST software (Watermark Numerical Computing, 2018) for calculation fidelity.

2.2.1. Sensitivities with respect to TCE discharge concentrations

Normalized sensitivity coefficients (X_{TCE}) were calculated on the basis of root mean squared errors (RMSE) between simulated (sim_i) and measured (obs_i) discharge concentrations as:

$$X_{TCE} = \left(\frac{|\partial RMSE| \; / \; RMSE(a)}{|\Delta a \; / \; a|} \right) / X_{TCE}^{max} \qquad (2)$$

where:

$$RMSE = \left[\frac{1}{N}\sum_{i=1}^{N}(sim_i - obs_i)^2\right]^{1/2} \quad (3)$$

$$\partial RMSE = RMSE(\Delta a) - RMSE(a) \quad (4)$$

a = base parameter; Δa = perturbed parameter – a; N = number of TCE effluent measurements. All sensitivity coefficients were normalized by maximum values (X_{TCE}^{max}) to provide a relative comparison metric of model sensitivities.

2.2.2. Sensitivities with respect to TOR

Provided with a cleanup concentration input, the VA model calculates the time required to reach the target value (e.g., contaminant MCL). Using the base parameter sets, which reflect detailed experimental conditions and initial source zone properties, TOR was calculated for both experiments setting target concentrations at C = 0.005 mg/L. Sensitivity coefficients normalized by maximum values (X_{TOR}^{max}) were calculated as:

$$X_{TOR} = \left(\frac{RMSE(\Delta a)}{|\Delta a \,/\, a|}\right) / X_{TOR}^{max} \quad (5)$$

3. Uncertainty Analysis

Source zone metrics controlling field-scale dissolution include DNAPL mass and distribution (Abriola et al., 2013). Uncertainties associated to both metrics can therefore propagate to model predictive uncertainties (Abriola, 2005; Tang, 2019). Prior (pre-history matching) parameter and predictive uncertainties can be informed by expert knowledge and/or by site characterization (e.g., DNAPL delineation, projected flow through area), whereas posterior (post-history matching) uncertainties may be reduced and quantified through history-matching of monitoring data. Because volume-averaging eliminates spatial parameter correlations, the prior uncertainty of mass transfer parameters was expressed through statistically uncorrelated uncertainty bounds (archived data file). All initial (mean) parameter values were inherited from Stewart et al. (2021).

Uncertainty bounds of characteristic dimensions (V_a) and mass (m_a) of DNAPL accumulations were designed so that $1\% < S_a^N < 60\%$ and $\sum V_a < V_S$ in both experiments; where DNAPL saturation (S_a^N) of the pore space (ϕ) is also a function of DNAPL density (ρ_n) as indicated by Equation 6. Per sensitivity results, upscaling parameters (V_S, U_0, and ϕ) were assumed well constrained by the monitoring scale and removed from predictive uncertainty evaluations. Uncertainty analyses were focused on m_a, V_a, α_T, and γ pertaining to each DNAPL accumulation. Linear and nonlinear uncertainty quantification methods were implemented to understand drivers of model uncertainties and bias emerging from data-driven conceptual assumptions.

$$S_a^N = \frac{m_a}{V_a \phi \rho_n} \quad (6)$$

3.1. Linear Analysis Methods

Model linearization expressed in Equation 7 is the primary assumption in first-order second-moment (FOSM) analysis (Doherty, 2015). Equation 7 indicates that a vector of measurements of system state **h** equals the action of the model **Z** on a vector of parameters **k** plus a vector of measurement noise **ε**. Prior model uncertainty was expressed by Equation 8 assuming a multi-gaussian probability density function (PDF), defined by mean parameter values $\underline{\mathbf{k}}$ and a diagonal covariance matrix C(**k**). Likewise, FOSM analysis assumes a multi-gaussian PDF of **ε** (Equation 9), defined by mean values of zero and a diagonal covariance matrix C(**ε**). Jacobian matrices **Z** were weighted by the inverse of the standard deviation (σ) of **ε**. The misfit between simulated (Stewart et al., 2021) and measured TCE concentrations was used to define **ε**, where $\sigma_{\varepsilon}^{-1}$ values were calculated with the PEST-based utility PWTADJ2 (Watermark Numerical Computing, 2018) as observations weights for FOSM analyses.

$$\mathbf{h} = \mathbf{Z}\mathbf{k} + \boldsymbol{\varepsilon} \tag{7}$$

$$\mathbf{k} \sim N[\underline{\mathbf{k}}, C(\mathbf{k})] \tag{8}$$

$$\boldsymbol{\varepsilon} \sim N[\mathbf{0}, C(\boldsymbol{\varepsilon})] \tag{9}$$

$$s = \mathbf{y}^t\mathbf{k} \tag{10}$$

$$\sigma_s^2 = \mathbf{y}^t C(\mathbf{k})\mathbf{y} \tag{11}$$

$$C'(\mathbf{k}) = C(\mathbf{k}) - C(\mathbf{k})\mathbf{Z}^t[\mathbf{Z}C(\mathbf{k})\mathbf{Z}^t + C(\boldsymbol{\varepsilon})]^{-1}\mathbf{Z}C(\mathbf{k}) \tag{12}$$

$$\sigma'^2_s = \mathbf{y}^t C'(\mathbf{k})\mathbf{y} \tag{13}$$

Linearization of a model prediction *s* (Equation 10) depends on a vector of sensitivities of *s* (TOR) with respect to **k**, where the prior variance of *s* (Equation 11) is obtained through covariance propagation (Doherty, 2015). The posterior parameter covariance matrix (Equation 12), obtained by history-matching conditioning, was used to estimate posterior TOR uncertainty variance (Equation 13). All parameters were log-transformed to reduce their nonlinearity with respect to model outputs. Linear analyses were performed with the utility programs GENLINPRED and PREDUNC (Watermark Numerical Computing, 2018) to understand how TCE monitoring profiles constrain source zone properties, and thereby, TOR uncertainties.

3.1.1. Relative parameter uncertainty variance (RUVR) reduction

This statistical metric was used to evaluate the ability of dissolved TCE concentrations to reduce the prior uncertainty variance (σ_i^2) of each parameter (*i*) encapsulated in C(**k**). Equation 14 defines this metric upon extracting posterior parameter uncertainty variances (σ'^2_i) from C'(**k**) as:

$$RUVR_i = 1 - \frac{\sigma'^2_i}{\sigma_i^2} \tag{14}$$

306 *3.1.2. Prior and posterior parameter contributions to predictive uncertainty*
307
308 The contribution of an individual parameter to the uncertainty of a prediction is defined as the fall
309 of predictive uncertainty resulting from acquiring perfect knowledge of the parameter (Doherty,
310 2015). Hence, individual parameters were systematically removed from FOSM calculations to
311 investigate their relative contributions to TOR uncertainty. Because history-matching information
312 may be shared between several model parameters, the posterior contribution of a parameter could
313 increase in relation to its prior contribution, indicating a correlation with another parameter
314 (Doherty, 2015). While sensitivity analyses were useful to examine relative model error incurred
315 by perturbing individual parameters, considering parameter correlations for TOR uncertainty
316 estimation allowed assessing the worth of HRSC over history-matching for constraining the
317 models.
318
319 *3.1.3. Data-Worth Analysis*
320
321 The ability of spatial or temporal data to reduce the uncertainty of model predictions defines its
322 worth (Doherty & Moore, 2020; Finsterle, 2005). The worth of individual measurements of TCE
323 concentrations was quantified to understand how monitoring profiles reduce TOR uncertainty.
324 Data-worth analyses were also tied to parameter RUVR, further elucidating upon the additional
325 benefit of HRSC for constraining remaining model uncertainties.
326
327 **3.2. Nonlinear Analysis Methods**
328
329 Posterior TOR uncertainties were quantified using the iterative ensemble smoother PESTPP-iES
330 (White et al., 2020). Multi-gaussian prior parameter PDFs were defined by uncertainty bounds
331 spanning ± 2σ from initial (mean = μ) values, representing 95% confidence intervals. PESTPP-
332 iES undertakes Monte-Carlo sampling of parameter uncertainty bounds generating model
333 realizations (ensembles) which are upgraded with the Gauss-Levenberg-Marquardt (GLM)
334 optimization algorithm. Rather than simply fitting simulation results to data, PESTPP-iES can
335 generate observation ensembles considering multi-gaussian PDFs of $\boldsymbol{\varepsilon}$ (White, 2018). Here, all
336 experimental TCE concentrations were assigned an observation weight value of 1 with $\sigma_{\boldsymbol{\varepsilon}} = 10$
337 mg/L, to simultaneously estimate model parameters and quantify the nonlinear uncertainty of TOR
338 in a stochastic manner. This approach was implemented to evaluate TOR uncertainties and bias
339 arising from source zone conceptual assumptions driven by data availability.
340
341 In practice, HRSC data may help constrain source zone architecture, and thereby model
342 conceptualizations. However, high predictive uncertainties may remain because of the inability to
343 directly measure DNAPL mass and S_a^N. The benefit of data assimilation for constraining model
344 uncertainties was investigated by estimating m_a, V_a, $\alpha_{T,a}$, and γ^a parameters in both experiments
345 from partial and complete monitoring profiles. The resulting source dissipation timeframes were
346 referred to as Posterior A (~13 days of monitoring), Posterior B (20 days) and Posterior All (26
347 days). Additionally, the heterogeneous experiment was conceptualized with 2 (2M), 3 (3M) and 4
348 (4M) DNAPL accumulations to examine TOR uncertainty and bias induced by history-matching
349 of the entire TCE monitoring profile. The 2M model included "mass 1" and lumped "mass 2" and
350 "mass 3" into a single accumulation (2M-3) based on the two release points, while the 3M model
351 included those 3 distinct DNAPL accumulations. The 4M model subdivided "mass 1" into 1A and

1B (Figure 2b). Except for the 2M Model, the 3M and 4M models included an enhanced dissolution parameter to represent flow channelization through the coarse lens in which "mass 3" was embedded. Following a variability range reported in the literature (Klenk & Grathwohl, 2002), the prior uncertainty bounds of α_T parameters were defined as $5e^{-4} < \alpha_T$ (m) $< 2e^{-3}$ in both experiments, except for 0 ($1e^{-15}$) $< \alpha_{T,3}$ (m) < 0.002 in the coarse sand lens of the heterogeneous experiment, where $\alpha_{T,3} = 0$ m provided the best match to measured TCE concentrations (Stewart et al., 2021).

4. Results and Discussion

4.1. Sensitivity Analysis

As shown in Figures 3 and 4, the greatest model sensitivities with respect to matching TCE concentrations (X_{TCE}) corresponded to the source zone area (Z_S and Y_S) orthogonal to the flow direction and groundwater velocity (U_0). The former accounts for any dilution in the monitoring scale, while the latter had a prominent impact on TOR in both experiments. The role of V_S and U_0 on scaling mass transfer processes emphasized the need to constrain them by the monitoring scale to avoid model errors induced by data assimilation. Figure 3 also indicated that the projected area facing flow (YZ_g) of the ganglia-dominated accumulation, rather than γ^g or ganglia mass, was responsible for peak aqueous-phase concentrations. Similarly, Figure 4 shows the projected area YZ_{1A} of the most upgradient, low-saturation accumulation 1A in a high-ranked position. These X_{TCE} results suggested that S_a^N parameters (V_a and $Mass$) of ganglia-dominated accumulations responsible for peak concentrations do not impact TOR when a pool-dominated accumulation is also present; yet their estimation via history-matching may be valuable for remedial designs. Conversely, sensitivity with respect to TOR (X_{TOR}) was dominated by DNAPL pool saturation (S_p^N) parameters, transverse dispersivity ($\alpha_{T,p}$), and depletion exponent (γ^p). The negligible X_{TCE} values of pool parameters suggested difficulty in estimating them from monitoring data alone and value in HRSC for refining characteristic parameters of the pool.

In contrast to negligible X_{TCE} values by pool parameters in the mixed experiment (Figure 3), X_{TCE} rankings of S_3^N parameters in the heterogeneous experiment (Figure 4) suggested that high-saturation DNAPL accumulations may not exclusively reflect pool fractions of source zones. Typically, the small cross-sectional areas available for dissolution by groundwater flow through DNAPL pools reduces their relative contribution to mass flux, compared to ganglia-dominated accumulations. However, as indicated in Figure 4, the morphology of DNAPL accumulation 3, controlled by flow-field heterogeneity, influenced both X_{TCE} and X_{TOR} rankings in the heterogeneous experiment. The predictive advantage of generalizing mass transfer processes irrespective of S_a^N (Equation 1) over upscaled models predicated on the GTP mass ratio, was further evidenced by a similar effect on X_{TCE} and X_{TOR} incurred by perturbing $\alpha_{T,3}$ (Figure 4). Conversely, the variability of other α_T parameters in both experiments only influenced X_{TOR}.

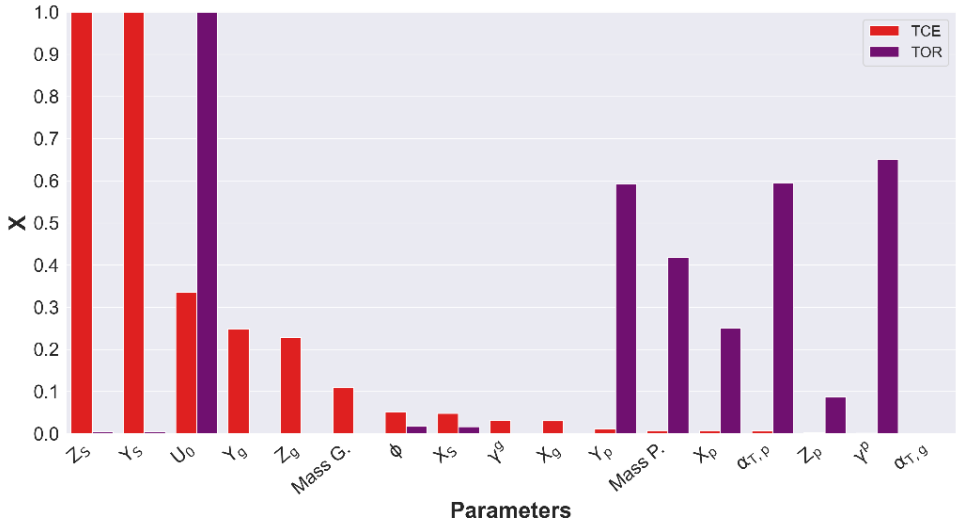

Figure 3. Sensitivity coefficients with respect to source discharge concentrations measured in the "mixed architecture" experiment and with respect to the simulated TOR.

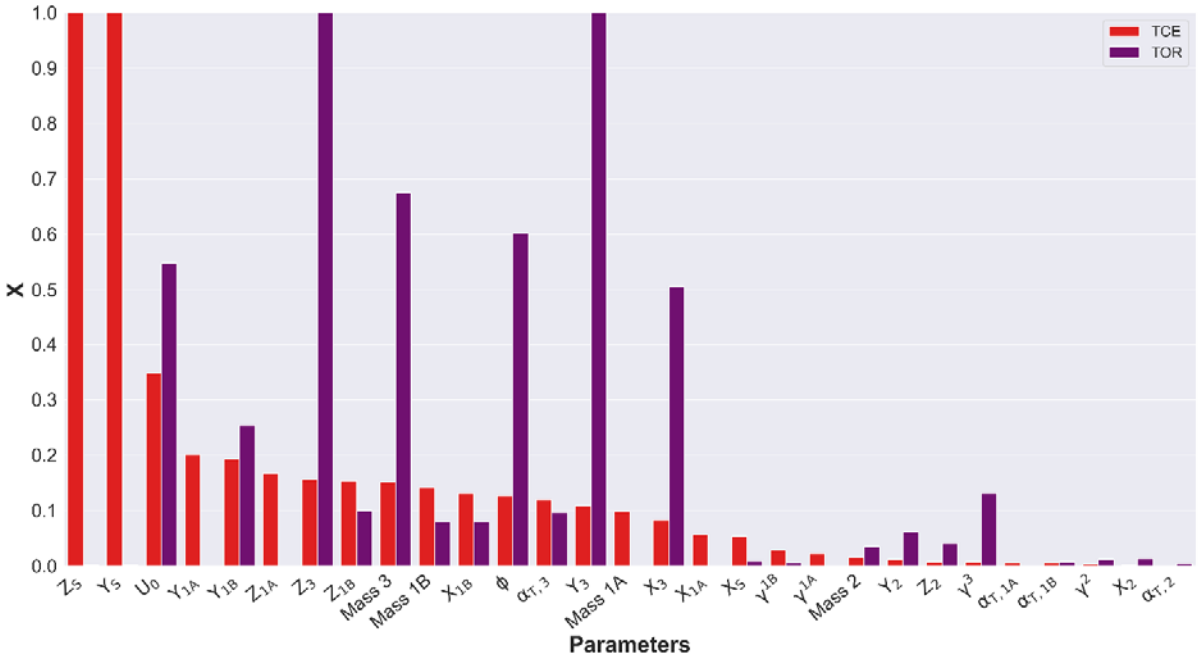

Figure 4. Sensitivity coefficients with respect to source discharge concentrations measured in the "heterogeneous architecture" experiment and with respect to the simulated TOR.

4.2. Linear Analysis

Prior (σ_{TOR}) and posterior (σ'_{TOR}) standard deviations of TOR uncertainty estimated with FOSM analysis and mean (μ_{TOR}) values for both experiments are presented in Table 2. Results shown were calculated using the complete TCE monitoring profiles. As indicated, history-matching significantly constrained prior TOR uncertainties despite low X_{TCE} values of TOR-sensitive parameters pertaining to high-S_a^N accumulations.

Table 2. Predictive uncertainty of mixed and heterogeneous experiments.

Experiment	μ_{TOR} (days)	σ_{TOR} (days)	σ'_{TOR} (days)
Mixed	27.9	19.8	8.6
Heterogeneous	28.6	20.5	1.7

4.2.1. Relative Parameter Uncertainty Variance Reduction

Figures 5 and 6 show the benefit of history-matching for reducing prior parameter uncertainties. Despite negligible X_{TCE} values corresponding to the pool mass and $\alpha_{T,p}$ of the mixed experiment (Figure 3), history-matching reduced the prior uncertainty of these parameters by ~70% and ~60%, respectively (Figure 5). The low uncertainty reduction of γ^p (Figure 5), to which TOR was sensitive (Figure 3), demonstrated the benefit of coupling upscaled modeling with stochastic analysis tools for predicting DNAPL longevity timeframes when mass transfer parameters remain unconstrained. In turn, sensitivity and FOSM analyses of the mixed experiment coincided in a low-ranked $\alpha_{T,g}$, suggesting that its prior (default) value of 0.001 m is reasonable for simulating dissolution of ganglia-dominated accumulations.

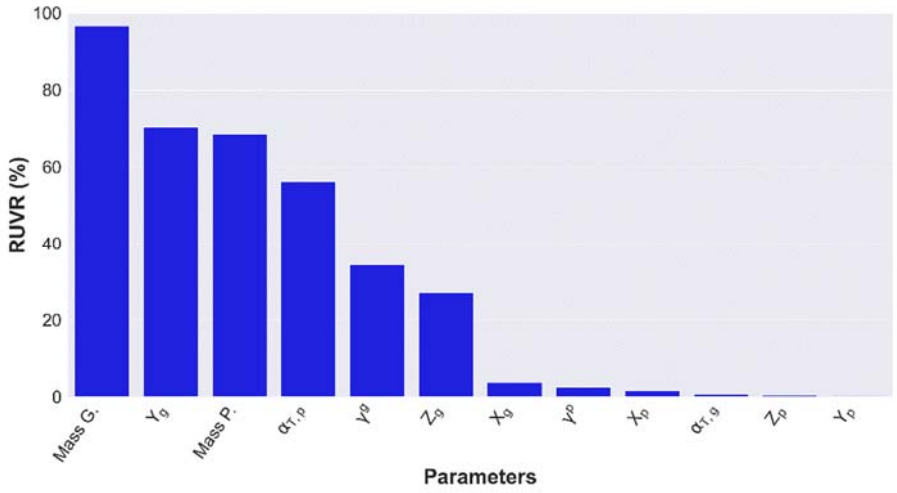

Figure 5. Relative uncertainty variance reduction of VA model parameters of mixed experiment.

Difficulties in reducing prior uncertainty of the γ parameters in the heterogeneous experiment are reflected in Figure 6. Yet the prior uncertainty of S_a^N parameters of DNAPL accumulations 1A (S_{1A}^N), 1B (S_{1B}^N), and 3 (S_3^N) was reduced by approximately more than 50%. The higher RUVR of S_3^N with respect to other S_a^N parameters was attributed to the sequential dissolution of upgradient DNAPL masses, allowing the tailing segment of the TCE monitoring profile to constrain the remaining source architecture (S_3^N). These results implied that modeling efforts supporting the characterization of sites with aged, pool-dominated source zones, may benefit from history-matching of monitoring profiles. However, situations with scarce monitoring data and significant uncertainties on S_a^N distributions along groundwater flow paths may warrant HRSC efforts. In turn, source characterization data such as DyeLIF and Hydraulic Profiling Tool (HPT) (Horst et al., 2018) can be leveraged for VA model parameterization, while FOSM analyses can help guide additional data collection efforts to constrain DNAPL dissolution trends.

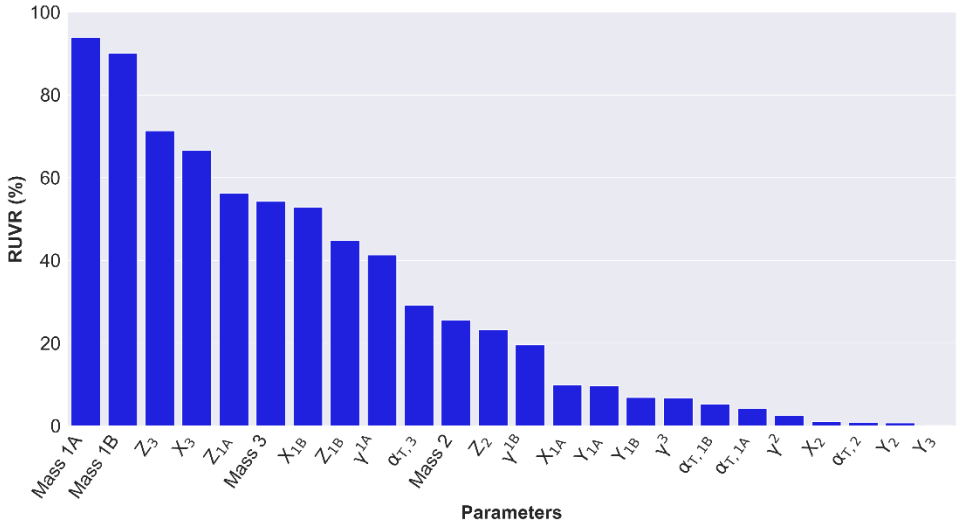

Figure 6. Relative uncertainty variance reduction of VA model parameters of heterogeneous experiment.

4.2.2. Prior and posterior parameter contributions to predictive uncertainty

As shown in Figure 7, FOSM analyses validated negligible X_{TOR} values caused by the ganglia parameters in the mixed experiment. Although the pool dispersive area (YX_p) and γ^p influenced X_{TOR} results (Figure 3), prior and posterior TOR uncertainties of the mixed experiment were clearly driven by the pool mass and $\alpha_{T,p}$ (Figure 7). Likewise, Figure 8 indicated that the primary drivers of prior TOR uncertainty in the heterogenous experiment were S_3^N, $\alpha_{T,3}$, and γ^3. Repeating FOSM calculations with uncertainty bounds defined as $0 < \alpha_T$ (m) < 0.01 for all DNAPL accumulations in the heterogeneous experiment did not alter the uncertainty rankings shown in Figure 8. Results of both experiments agreed on the significance of dispersive mass transfer (α_T) from high-saturation DNAPL accumulations in regulating TOR. However, the accurate replication of the heterogenous source dissolution trend with $\alpha_{T,3} = 0$ m was attributed to the contrast in grain sizes, limiting dispersion from the coarse-grained lenticular zone into the finer surrounding sands despite high $S_{3,0}^N$ values.

Unlike the empirical mass depletion exponent γ^3, α_T may be directly measured at contaminated sites to directly constrain mass transfer uncertainties. Examples of field methods include push-pull tracer tests, borehole and HPT logging, and discrete groundwater sampling with direct push technology (DPT). These data may be interpreted with 2D analytical modeling (Huang et al., 2010), grain-size correlations with soil hydraulic conductivity and α_T (Carey et al., 2018), and spatial moment analysis (Rockhold et al., 2016), respectively. Nonetheless, the α_T component of DNAPL dissolution expressed in Equation 1 should not be confused with plume-scale macrodispersion. While dispersivity at the source-zone and plume scales is driven by mechanical or hydrodynamic mixing along tortuous flow paths (Molz, 2015), coupling a VA model of DNAPL dissolution with a downgradient contaminant plume model may require two different α_T values based upon site-specific conditions. Several studies have demonstrated the relationship between soil grain size and α_T (Carey et al., 2018), concurring with its role on DNAPL mass transfer (Figures 7 and 8). This is in contrast to Gilland-Sherwood mass transfer correlations which rely upon aqueous-phase transport models for the contribution of α_T to DNAPL dissolution (Yang et al., 2019).

As indicated in Figure 8, the primary driver of posterior TOR uncertainty, γ^3, reflected its role in regulating source discharge concentrations over several orders of magnitude. While a lack of extensive groundwater monitoring at contaminated sites could limit γ constraining via history-matching, S_a^N and flow-field heterogeneities may also pose additional uncertainties on mass transfer assumptions. In this case, TCE dissolution tailing, primarily regulated by S_3^N, was also modulated by flow channelization in the coarse sand lens (Figure 2b). Transient reductions in DNAPL interfacial areas, which limit mass transfer rates through the γ parameter, were obfuscated by a local increase in U_0 and k_r in the heterogeneous experiment (Stewart et al., 2021). Although the level of characterization detail available for the flow-cell experiment would not be available at field sites, VA modeling provides an efficient means to evaluate conceptual assumptions of system heterogeneities and quantify mass transfer uncertainties. The prior uncertainty rankings of S_{1B}^N and S_3^N parameters (Figure 8) emphasized the level of effort for DNAPL delineation required for adequate model parameterization.

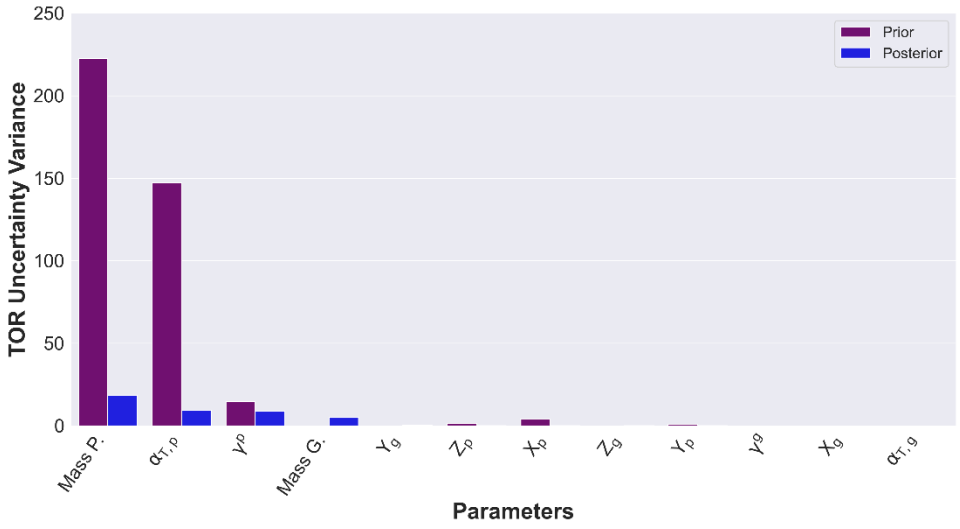

Figure 7. Prior and posterior parameter contributions to TOR uncertainty in the mixed experiment.

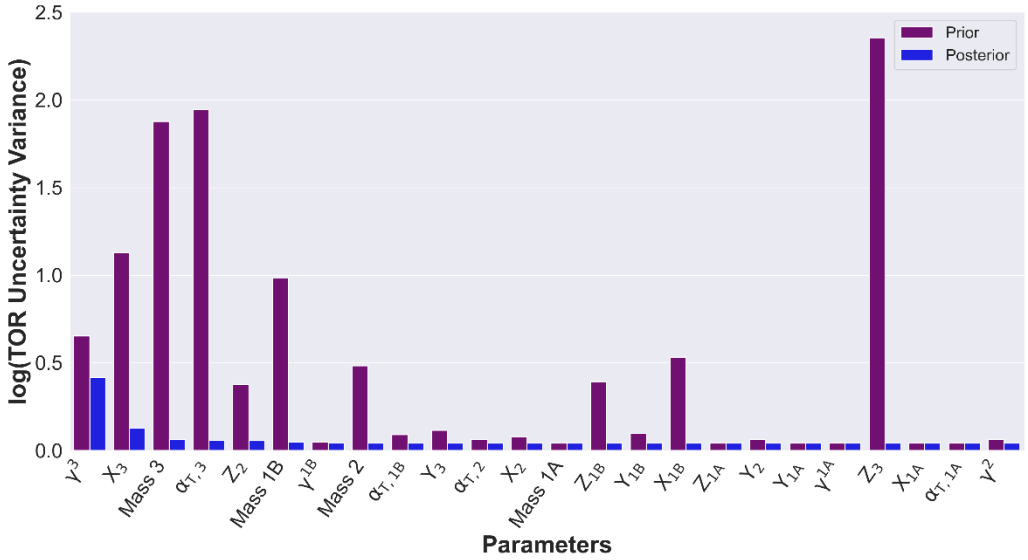

Figure 8. Prior and posterior parameter contributions to TOR uncertainty of heterogeneous experiment.

4.2.3. Data-Worth Analysis

The worth of TCE monitoring profiles of the mixed and heterogeneous experiments is shown in Figures 9 and 10, respectively. Both figures express data worth on the Y-axis as a percent reduction and increase of σ_{TOR} and σ'_{TOR} (see Table 2), respectively, by individual monitoring measurements. Figures 9a and 10a indicate the worth of individual measurements for constraining prior TOR uncertainty (σ_{TOR}), whereas Figures 9b and 10b depict increases in posterior (constrained) TOR uncertainty (σ'_{TOR}) caused by data removal. As shown in Figure 10, a tendency of increasing data worth in the mixed experiment started at point C, when the pool mass transfer area ($A_{p,xy}$) was sufficiently reduced to onset dissolution tailing. Similar prior and posterior data-worth trends in the mixed experiment suggested that peak concentrations emanating from ganglia-dominated accumulations do not constrain TOR. In turn, the RUVR of pool mass (~70%) and $\alpha_{T,p}$ (~60%) controlling TOR uncertainty was attributed to TCE monitoring after point C (Figure 9), highlighting the benefit of history-matching for characterizing sites with aged source zones and simple architectures. In these experiments, point C represents a rough mid-point for the DNAPL TOR despite an 80% reduction in the total DNAPL mass.

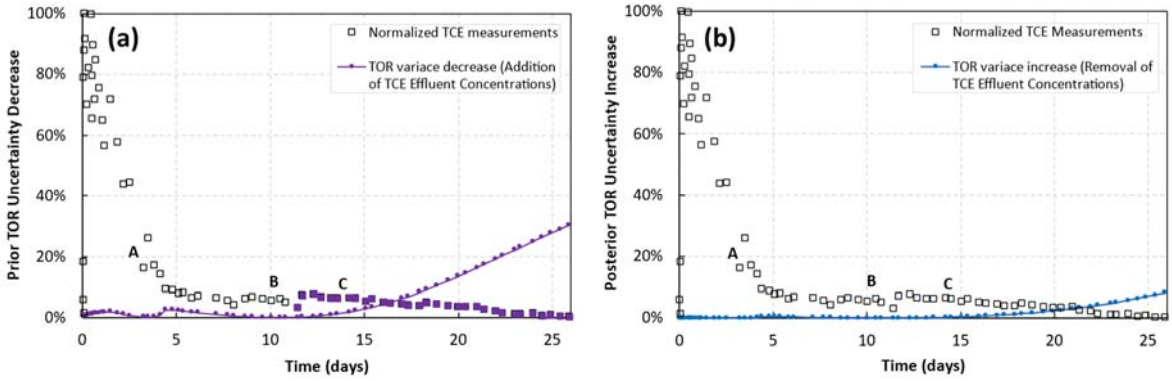

Figure 9. Worth of monitoring data for constraining TOR uncertainty of the mixed experiment shown in Figure 2a: a) Decrease in prior uncertainty with addition of individual TCE concentrations. The filled data points highlight the

greatest information content for reducing prior TOR uncertainty. b) Increases in posterior uncertainty with data removal. Points A, B, C show DNAPL depletion images measured by DiFilippo et al. (2010).

Figure 10a shows the worth of breakthrough inflection points along the TCE monitoring curve of the heterogeneous experiment for constraining σ_{TOR}. The first peak in the σ_{TOR} decrease curve coincided with point A, indicating the onset of rapid dissolution of DNAPL mass accumulation 1b after mass 1a was completely dissolved. The second peak of σ_{TOR} reduction occurred during a slight increase in TCE concentrations, reflecting an increased k_r through mass 2 after mass 1B was dissolved. The final peaks of σ_{TOR} reduction (Figure 10a) and σ'_{TOR} increase (Figure 10b) coincided with the final stage of DNAPL dissolution associated to mass 3. These results highlighted disadvantages of predicting future system behavior from limited monitoring profiles, corresponding to situations where remaining source architectures and heterogeneities have not yet been reflected in historical dissolution trends.

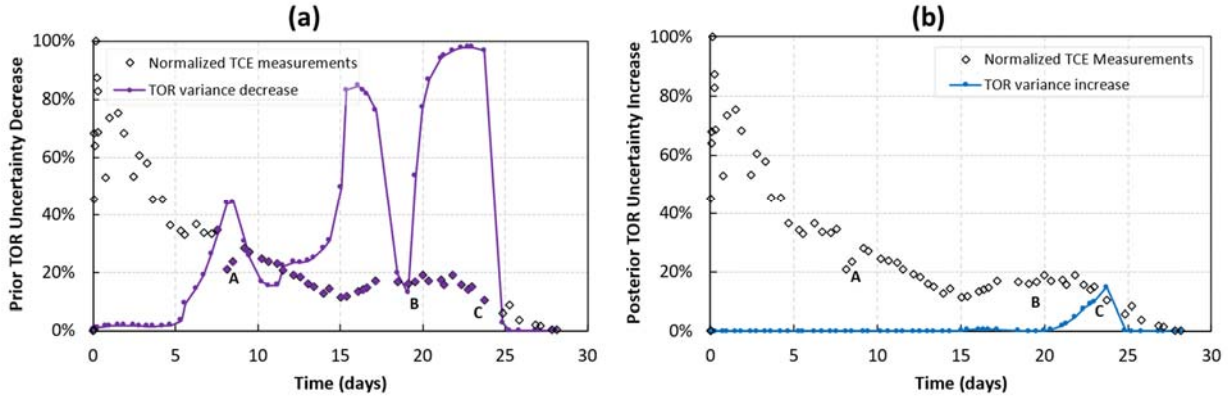

Figure 10. Worth of TCE dissolution measurements for reducing TOR uncertainty of the heterogeneous experiment shown in Figure 2b: a) Decrease in prior uncertainty with addition of individual history-matching constraints. The filled data points highlight the greatest information content for reducing prior TOR uncertainty. b) Increases in posterior uncertainty with data loss. Points A, B, C show the DNAPL depletion measured by DiFilippo et al. (2010).

4.3. Nonlinear Uncertainty Analysis

Figure 11 indicates that all prior and posterior source dissipation timeframes of the mixed experiment included the "true TOR" (μ_{TOR} = 27.9 days). All posterior analyses underestimated the initial DNAPL mass in the mixed experiment by ~11%, yet the known value of 17.2 g was included within 95% confidence limits (results not shown). The average estimated S_g^N and S_p^N values were 4% and 40%, respectively, consistent with initial experimental conditions (Figure 2a). Prior and posterior TOR uncertainties in Figure 11 demonstrated the utility of VA modeling for estimating unbiased depletion timeframes a priori, by leveraging DNAPL-delineation or limited monitoring data pertaining to source zones with relatively simple architectures and flow conditions.

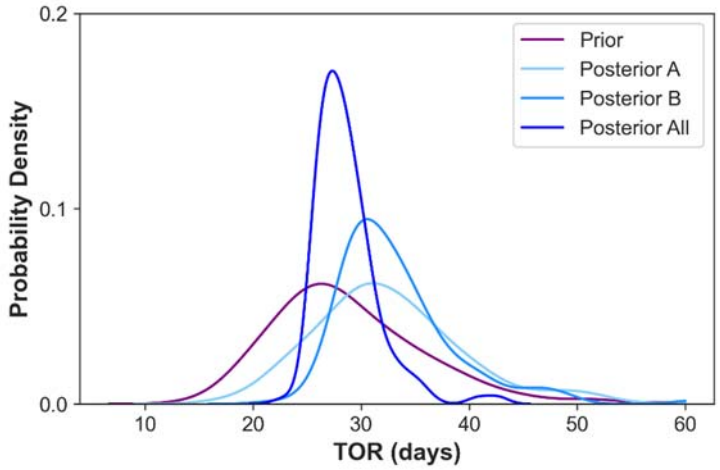

Figure 11. Prior and posterior TOR PDFs of mixed experiment. Posterior A and B were estimated by history-matching TCE concentrations through day 11.7 and 20 (Figure 10), respectively.

The stochastic optimization of the heterogeneous experiment models underestimated initial DNAPL mass by ~7%, with 95% confidence limits encompassing the injected amount of 20.4 g (results not shown). As illustrated in Figure 12, posterior S_a^N uncertainties reflected the averaging by model resolutions required to history-match the complete TCE dissolution profile and quantify TOR uncertainty (Figure 13). Figure 13 shows all posterior TOR PDFs encompassing the "true" TOR of 28.6 days, emphasizing the worth of final DNAPL dissolution stages for constraining TOR with various model resolutions. However, the 2M and 3M models required removing peak TCE concentrations from day 0 through day 9 (Figure 14). Not doing so did not impact the accuracy of estimated DNAPL mass, but resulted in an artificially low initial S_I^N of lumped mass 1 from inadequate parameterization complexity (results not shown). Sufficient source architecture parameters are thereby necessary to assimilate complex dissolution profiles to avoid misleading injection-based remedial designs.

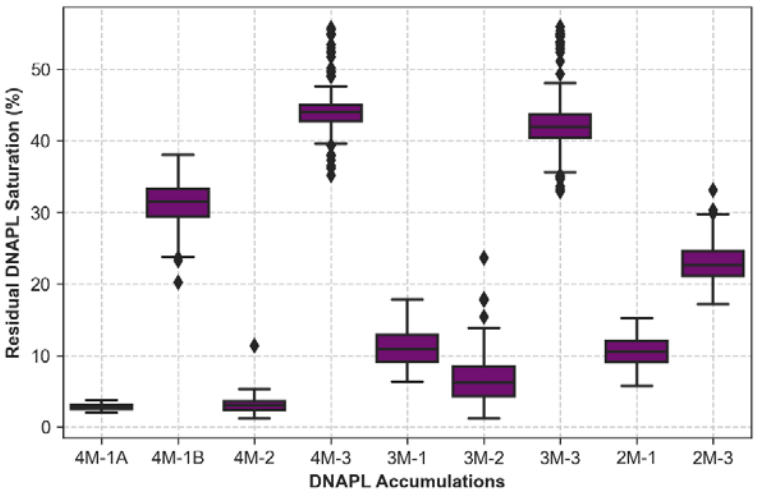

Figure 12. Posterior DNAPL saturation distributions of each DNAPL accumulation in the 4M, 3M, and 2M VA models of the heterogeneous experiment.

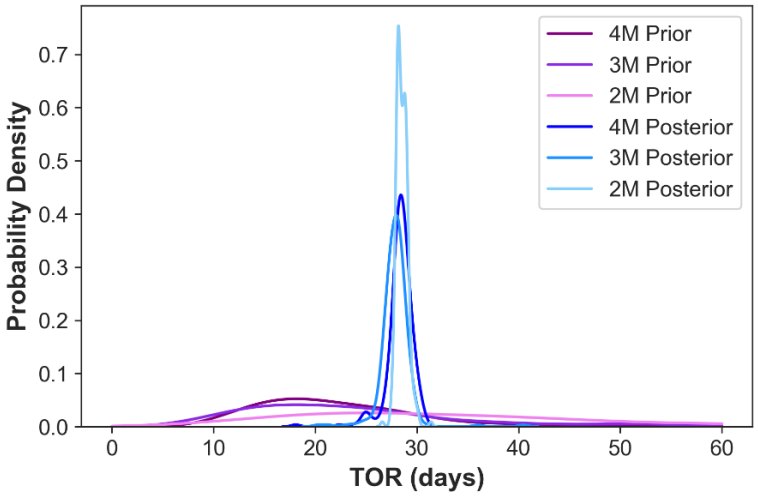

Figure 13. Prior and posterior TOR PDFs of the heterogeneous experiment conceptualized by 2, 3, and 4 DNAPL accumulations.

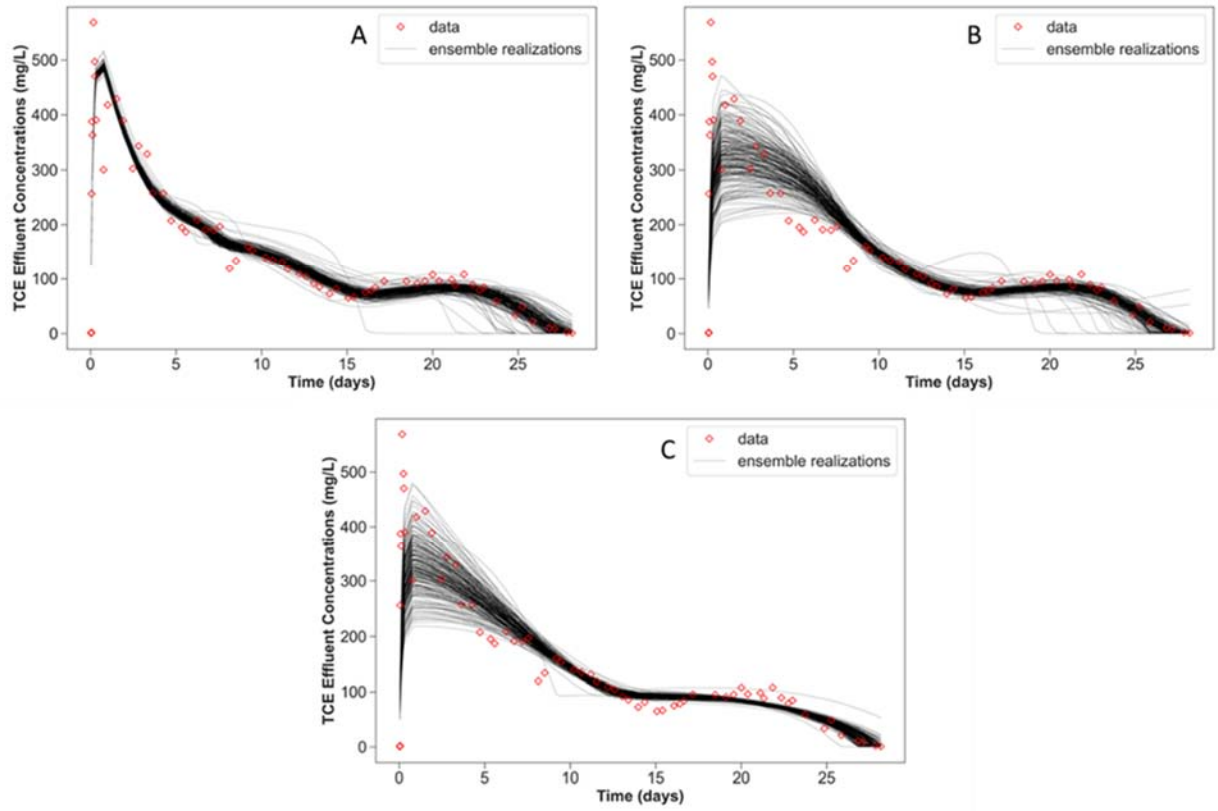

Figure 14. Posterior model ensembles of the heterogeneous experiment corresponding to (A) 4, (B) 3, and (C) 2 DNAPL accumulations.

Figure 15 shows prior predictive PDFs approximated with S_a^N constraints assuming availability of HRSC data to inform model parameters. The prior S_a^N constraints for this analysis correspond Figure 12. Despite low probability densities, all prior PDFs encompassed the μ_{TOR} = 28.6 days, suggesting that even a low-resolution model (2M) accounting for S_a^N distributions along the flow

path can predict unbiased source dissipation timeframes. However, Figure 15 depicts biased posterior 2M PDFs tending to exclude μ_{TOR} resulting from monitoring data assimilation with inadequate parameterization complexity. Unlike 2M, 3M included an adjustable "dissolution enhancement factor" representing increased velocity through the coarse lens. Omitting that parameter from the 3M model (fixing it at a value of 1) did not impact σ'_{TOR} estimated from the entire TCE profile (Figures 14 and 15). However, σ_{TOR} and σ'_{TOR} estimated from partial TCE profiles were also overestimated (excluding μ_{TOR}) and the nonmonotonic increase in TCE concentrations from day 15 through day 20 could not be reproduced (results not shown). Hence, the unbiased Posterior A and B results of the 3M model, shown in Figure 15, suggested that in addition to adequate representation of DNAPL distribution along the local flow path, parameterization of flow field heterogeneity is also necessary to avoid biasing model estimates through history-matching of multistage and nonmonotonic dissolution profiles.

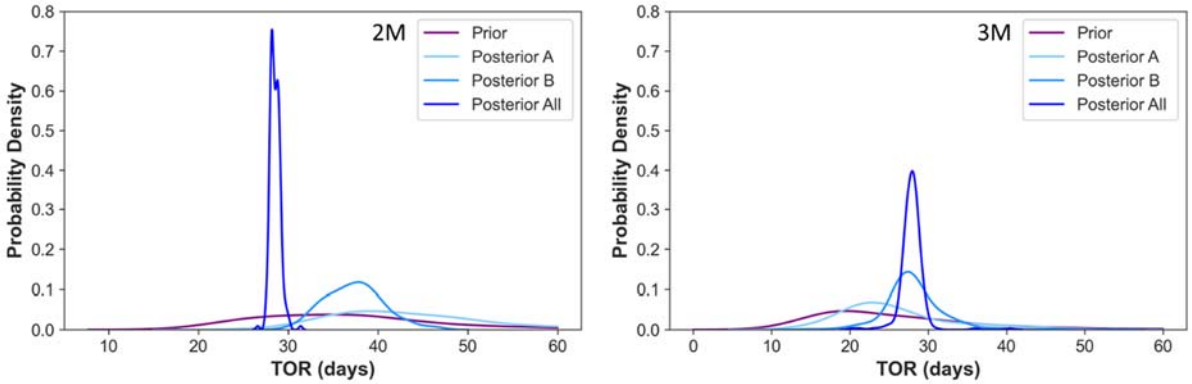

Figure 15. Probability density functions of TOR approximated with 2M and 3M models of heterogeneous experiment. Posterior A and B PDFs were estimated from partial TCE monitoring profiles through day 14 and 20, respectively.

5. Conclusions

This work demonstrated a practical approach for estimating DNAPL dissolution timeframes coupling upscaled modeling with uncertainty analysis methods. Assimilation of monitoring data may induce model predictive bias without sufficient parameterization complexity representing the DNAPL source, including sequential dissolution of DNAPL accumulations distributed along the flow path. In both experiments, saturation parameters and transverse dispersion of pool-dominated DNAPL accumulations controlled the source zone longevity, and were constrained by tailing of final dissolution stages despite their negligible sensitivity with respect to measured effluent concentrations. Because the VA model provides TOR as a direct output, FOSM analyses can be used to guide site characterization efforts to constrain prior, or remaining posterior parameter uncertainties responsible for predictive TOR and mass discharge/flux uncertainties. As demonstrated with the heterogeneous source zone experiment, field mapping of aquifer hydraulics, and/or estimation of source zone architecture using physically-based inversion methods can be leveraged to refine site conceptual assumptions encapsulated in VA model parameters. This includes direct constraining of transverse vertical dispersivity at the source zone scale, regardless of DNAPL saturation, differentiating its contribution to DNAPL dissolution from macrodispersion at the contaminant plume scale.

Local groundwater velocity and source zone dimensions had a prominent impact on mass discharge and DNAPL persistence because of their scaling role on mass transfer processes. Hence, these parameters require constraining by monitoring and site characterization scales, promoting adequate dilution and flow bypassing effects on DNAPL dissolution. Conversely, saturation parameters of ganglia-dominated DNAPL accumulations, which may not be directly measured at field sites, did not impact source longevity timeframes when pools were present. Yet their influence on peak discharge concentrations justifies their parameterization to avoid erroneous estimates of DNAPL saturation distributions and mass discharge rates. Although accurately simulating mass discharge was possible with increased resolution of source zone architecture, exclusive designations of ganglia and pool fractions of DNAPL may be inadequate for mass transfer modeling. The high-saturation DNAPL accumulation embedded in the coarse sand lens of the heterogeneous experiment, controlled the source zone longevity without dispersive mass transfer. Moreover, lumping the downgradient saturations and ignoring flow field heterogeneity, biased lifespan estimates of the heterogenous source zone and degraded the replication of nonmonotonic DNAPL dissolution tailing. While this level of characterization detail may not be available for contaminated sites, upscaled modeling and stochastic uncertainty analyses of site conceptual assumptions can support risk-based decision making through data assimilation and predictive hypothesis testing with a physical mass transfer basis.

Acknowledgements

This study was 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

A copy of the VA model executable and input instructions can be requested from Praxis Environmental Tech., Inc. at https://www.praxis-enviro.com/contact. The SENSAN, PWTADJ2, PREDUNC, GENLINPRED, and PEST software utilities used for sensitivity and linear uncertainty analyses are available at https://pesthomepage.org/programs. The PESTPP-iES software used for ensemble-based parameter estimation and nonlinear uncertainty analyses 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). Except for the data-worth results figures, figures in the results and discussion section 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. Data is supplied in an excel file for peer review purposes and will be archived in an online repository maintained by Virginia Tech with a unique DOI number.

6. References

Abriola, L. M. (2005). Guest Editorial: Contaminant Source Zones: Remediation or Perpetual Stewardship? *Environmental Health Perspectives, 113*(7), A438-A439. https://doi.org/10.1289/ehp.113-a438

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

Carey, G. R., McBean, E. A., & Feenstra, S. (2018). Estimating transverse dispersivity based on hydraulic conductivity. *Environmental Technology & Innovation, 10*, 36-45. https://doi.org/10.1016/j.eti.2018.01.008

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

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

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

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

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

693 DiFilippo, E. L., Carroll, K. C., & Brusseau, M. (2010). Impact of organic-liquid distribution and
694 flow heterogeneity on reductions in mass flux. *Journal of Contaminant Hydrology, 115*(1-
695 4), 14-25. https://doi.org/10.1016/j.jconhyd.2010.03.002

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

698 Doherty, J., & Moore, C. (2020). Decision Support Modeling: Data Assimilation, Uncertainty
699 Quantification, and Strategic Abstraction. *Groundwater, 58*(3), 327-337.
700 https://doi.org/10.1111/gwat.12969

701 Eniarson, M., Fure, A., St. Germain, R., Chapman, S., & Parker, B. (2018). DyeLIF™: A New
702 Direct-Push Laser-Induced Fluorescence Sensor System for Chlorinated Solvent DNAPL
703 and Other Non-Naturally Fluorescing NAPLs. *Groundwater Monitoring & Remediation*,
704 28-42. https://doi.org/10.1111/gwmr.12296

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

708 Finsterle, S. (2015). Practical notes on local data-worth analysis. *Water Resources Research*.
709 https://doi.org/10.1002/2015WR017445

710 Frind, E. O., Molson, J. W., & Schirmer, M. (1999). Dissolution and mass transfer of multiple
711 organics under field conditions: The Borden emplaced source. *Water Resources Research,*
712 *35*(3), 683-694. https://doi.org/10.1029/1998WR900064

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

716 Horst, J., Welty, N., Stuetzle, R., Wenzel, R., & Germain, R. (2018). Fluorescent dyes: A new
717 weapon for conquering DNAPL characterization. *Groundwater Monitoring &*
718 *Remediation, 38*(1), 19-25. https://doi.org/10.1111/gwmr.12261

719 Huang, J., Christ, J. A., & Goltz, M. N. (2010). Analytical solutions for efficient interpretation of
720 single-well push-pull tracer tests. *Water Resources Research*.
721 https://doi.org/10.1029/2008WR007647

722 ITRC (Interstate Technology & Regulatory Council). (2010). *Use and Measurement of Mass Flux*
723 *and Mass Discharge.* Washington, D.C.: Interstate Technology & Regulatory Council,
724 Integrated DNAPL Site Strategy Team. Retrieved from www.itrcweb.org

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

729 Kang, X., Kokkinaki, A., Power, C., Kitandis, P. K., Shi, X., Duan, L., . . . Wu, J. (2021b).
730 Integrating deep learning-based data assimilation and hydrogeophysical data for improved
731 monitoring of DNAPL source zones during remediation. *Journal of Hydrology, 601*,
732 126655. https://doi.org/10.1016/j.jhydrol.2021.126655

733 Kang, X., Kokkinaki, A., Shi, X., Yoon, H., Lee, J., Kitandis, P. K., & Wu, J. (2022). Integration
734 of Deep Learning-Based Inversion and Upscaled Mass-Transfer Model for DNAPL Mass-
735 Discharge Estimation and Uncertainty Assessment. *Water Resources Research, 58*(10).
736 https://doi.org/10.1029/2022WR033277

737 Klenk, I. D., & Grathwohl, P. (2002). Transverse vertical dispersion in groundwater and the
738 capillary fringe. *Journal of Contaminant Hydrology, 58*(1–2), 111-128.
739 https://doi.org/10.1016/S0169-7722(02)00011-6

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

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

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

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

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

755 Lemke, L. D., & Abriola, L. M. (2006). Modeling dense nonaqueous phase liquid mass removal
756 in nonuniform formations: Linking source-zone architecture and system response.
757 *Geosphere, 2*(2), 74-82. https://doi.org/10.1130/GES00025.1

758 Marble, J. C., DiFilippo, E. L., Zhang, Z., Tick, G. R., & Brusseau, M. L. (2008). Application of
759 a lumped-process mathematical model to dissolution of non-uniformly distributed
760 immiscible liquid in heterogeneous porous media. *Journal of Contaminant Hydrology,*
761 *100*, 1-10. https://doi.org/10.1016/j.jconhyd.2008.04.003

762 Miller, C. T., Christakos, G., Imhoff, P. T., McBride, J. F., & Pedit, J. A. (1998). Multiphase flow
763 and transport modeling in heterogeneous porous media: challenges and approaches.
764 *Advances in Water Resources, 21*(2), 77-120. https://doi.org/10.1016/S0309-
765 1708(96)00036-X

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

769 Molz, F. (2015). Advection, Dispersion, and Confusion. *Ground Water, 53*(3), 348-353.
770 https://doi.org/10.1111/gwat.12338

771 Moore, C., & Doherty, J. (2005). Role of the calibration process in reducing model predictive
772 error. *Water Resources Research*. https://doi.org/10.1029/2004WR003501

773 Parker, J. C., & Park, E. (2004). Modeling field-scale dense nonaqueous phase liquid dissolution
774 kinetics in heterogeneous aquifers. *Water Resources Research, 2004*.
775 https://doi.org/10.1029/2003WR002807

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

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

783 Rockhold, M., Zhang, Z., & Bott, Y.-J. (2016). *Scale-Dependent Solute Dispersion in Variably*
784 *Saturated Porous Media.* Richland, WA: Pacific Northwest National Laboratory.

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

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

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

794 Tang, T. (2019). *An Adjoint-Sensitivity-Analysis Based Mathematical Framework: DNAPL Source*
795 *Zone Characterization, Uncertainty Quantification, and Sampling Strategy Design*
796 (Doctoral dissertation). Civil and Environmental Engineering. Ann Arbor, MI: Tufts
797 University. Retrieved from ProQuest (Access provided by University Libraries through
798 Virginia Tech)

799 Watermark Numerical Computing. (2018). *Model-Independent Parameter Estimation. User*
800 *Manual Part II: PEST Utility Suport Software.*

801 White, J. T. (2018). A model-independent iterative ensemble smoother for efficient history-
802 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., 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