

Predicting remediation efficiency of LNAPLs using surrogate polynomial chaos expansion model and global sensitivity analysis

Taehoon Kim^a, Weon Shik Han^{a,*}, Jize Piao^a, Peter K. Kang^{b,c}, Jehyun Shin^d

^a Department of Earth System Sciences, Yonsei University, Seoul, Republic of Korea

^b Department of Earth and Environmental Sciences, University of Minnesota, Minneapolis, MN, USA

^c Saint Anthony Falls Laboratory, University of Minnesota, Minneapolis, MN, USA

^d Groundwater Environment Research Center, Korea Institute of Geoscience and Mineral Resources, Daejeon, Republic of Korea

ARTICLE INFO

Keywords:

LNAPLs remediation
Removal efficiency
Polynomial chaos expansion
Sobol sensitivity analysis
Multi-phase modeling
Monte Carlo prediction

ABSTRACT

A multi-phase and multi-component numerical simulator assessed the removal efficiencies of light non-aqueous phase liquid (LNAPL) consisting of benzene, toluene, ethylbenzene, and xylene-p. Scenarios of the LNAPL-spilling, natural distribution, and remediation stages were designed in the full-physics numerical modeling. For LNAPLs remediation, a multi-phase extraction (MPE) and a steam injection technique were employed. The removal efficiencies of LNAPLs were computed by systematically varying 6 factors that determine the configuration of the remediation wells. Then, surrogate polynomial chaos expansion (PCE) models mathematically predicting the removal efficiencies were developed through 600 training datasets representing 4 scenario cases; different permeability and the location of SI well were considered in the scenario cases. The PCE models were utilized for Sobol global sensitivity analysis and stochastic Monte Carlo prediction. As a result, the depth of the MPE well was identified as the most significant factor in determining the removal efficiency of the LNAPLs. The removal efficiency was maximized when the MPE well was positioned 1.5 m below the groundwater table. Additionally, the contributions of influencing factors were significantly changed by the field permeability. This study proposed a general framework that efficiently predicts LNAPLs remediation efficiency and identifies key influencing factors by combining advanced numerical modeling, PCE-based surrogate modeling, and sensitivity analyses.

1. Introduction

Light non-aqueous phase liquids (LNAPLs) are typical contaminants that originate from various anthropogenic sources such as oil storage tanks, gas stations, transportation, and military camps (Essaid et al., 2015; Huntley and Beckett, 2002; Oostrom et al., 2006). LNAPLs lighter than water float above the groundwater table and extensively contaminate both soil and groundwater zones, when they spill into the subsurface. LNAPLs contain deleterious components including benzene, toluene, ethylbenzene, and xylene-p (BTEX). These components can be volatilized into the unsaturated zone and concurrently dissolved into the groundwater while LNAPLs drifting above the groundwater table. Therefore, the thermodynamic properties of LNAPLs that elucidate comprehensive phase transfer (i.e., volatilization, condensation, and dissolution) should properly be assessed. However, the assessment is challenging due to discrepancies in thermodynamical behaviors of

multiple components (BTEX) within LNAPLs (Adenekan et al., 1993). Additionally, migration of LNAPLs reflects multi-phase transport within porous media, representing the coexistence of LNAPLs, air, and water. In this complex transport system, both relative permeability and capillary pressure must be evaluated (Brennen and Brennen, 2005; Oliveira and Demond, 2003).

The complex dynamics of LNAPLs transport in the subsurface can be elucidated by conducting numerical simulations. Over the past two decades, many studies have attempted to predict the migration of spilled LNAPLs via numerical simulations (Gupta et al., 2019; Kim and Corapcioglu, 2003; Qi et al., 2020; Schroth et al., 1998; Wipfler et al., 2004). Although such studies successfully investigated the multi-phase behavior of LNAPLs, their approach limited LNAPLs as a single component while LNAPLs are generally observed as a mixture of several components (Essaid et al., 2015). Recently, Yang et al. (2017) simulated multi-component LNAPLs transport to evaluate the removal ratio of

* Corresponding author.

E-mail address: hanw@yonsei.ac.kr (W.S. Han).

<https://doi.org/10.1016/j.advwatres.2022.104179>

Received 1 December 2021; Received in revised form 3 March 2022; Accepted 23 March 2022

Available online 25 March 2022

0309-1708/© 2022 Elsevier Ltd. All rights reserved.

BTEX through soil vapor extraction. The multi-phase multi-component LNAPLs simulation could be further enhanced by accounting for hysteresis effects and 3-D heterogeneity in permeability (Sookhak Lari et al., 2018). The results aided in simulating LNAPL transport realistically and improved the assessment of LNAPLs removal efficiency. However, those full-physics numerical modelings and subsequent analyses are not always viable; such complex numerical analyses require enormous CPU time and memory spaces suited to supercomputing systems, which may not be available to all users.

Data-driven surrogate modeling techniques have been suggested as alternative options, to compensate for the difficulties caused by complexity in numerical models and to manage consecutive post-analyses with heavy iterative computations (Asher et al., 2015; Razavi et al., 2012). The surrogate modeling technique generates an empirical formula from several pairs of given input and output datasets predicted from complex numerical models. The application of the formula is limited to the designed model domain (or site-specific) represented by the chosen numerical simulator. Furthermore, complex numerical models, which is a basis of surrogate models, may not provide a perfect description of physics as well as subsurface heterogeneity. Despite limitations addressed above, the surrogate modeling technique has been in the limelight as decision support in the field of hydrogeology due to its ability to represent the dominant features in the subsurface with cheaper computational costs (Asher et al., 2015). Since the verified surrogate models allow for the prediction of the outputs without requiring heavy numerical simulations, the demanding time for the computation is significantly reduced. A key advantage from the reduction in the computational time is the availability of post-analyses using stochastic methods. In recent years, the stochastic post-analyses based on the surrogate model have been used to analyze seawater intrusion (Rajabi et al., 2015), geologic CO₂ sequestration (Guyant et al., 2016; Jia et al., 2016; Piao et al., 2020), and shallow groundwater contamination (Ciriello and de Barros, 2020; Xing et al., 2019; Zhao et al., 2020). As seen, although the applications for the surrogate modeling techniques have expanded, limited surrogate modeling and post-analyses have been

attempted in the field of LNAPL transport and its remediation.

In this study, a state-of-the-art technique of surrogate modeling, the polynomial chaos expansion (PCE), was employed to link between the full-physics numerical LNAPL model and stochastic post-analyses. After developing the surrogate PCE model, Sobol global sensitivity analysis was conducted to rank important factors that affect multi-phase and multi-component LNAPL remediation. In addition, Monte Carlo analysis was conducted to comprehensively investigate the effects of each factor on remediation efficiency. Eventually, this study proposes an efficient framework that quantitatively investigates the influencing factors, and can be easily extended for optimizing various LNAPLs remediation scenarios.

2. Conceptual model

A 2D full-physics numerical model was developed to investigate the dynamics of LNAPLs in unsaturated and saturated zones. As shown in Fig. 1a, it was assumed that the LNAPLs, a mixture of BTEX, were spilled from a designated point source. The spilled LNAPL plume migrated downward through the unsaturated zone until reaching the groundwater table. Due to buoyancy force, the LNAPLs floated above the groundwater table. In the capillary fringe zone where LNAPLs, air, and groundwater coexist, LNAPL dynamics were governed by gravity and capillary pressure (Bear et al., 1996; Mayer, 2005) (Fig. 1b). Here, LNAPLs contemporaneously volatilized to the unsaturated zone as highly volatile components; hence BTEX components are also referred to as volatile organic compounds (VOCs) (Fig. 1c) (Kim and Corapcioglu, 2003; Russell, 1995). The amounts and rates of volatilized BTEX components are determined based on the partitioning coefficient and saturated vapor pressures (Mendoza and Frind, 1990). In addition, at the bottom of the LNAPL plume where the groundwater is contacted, some of the LNAPLs dissolve into groundwater (Fig. 1d); the amounts of partially dissolved LNAPLs are determined by the solubilities of each BTEX component. Such addressed interactions, including chemical dissolution and thermophysical phase transfer, affect the fate of LNAPL

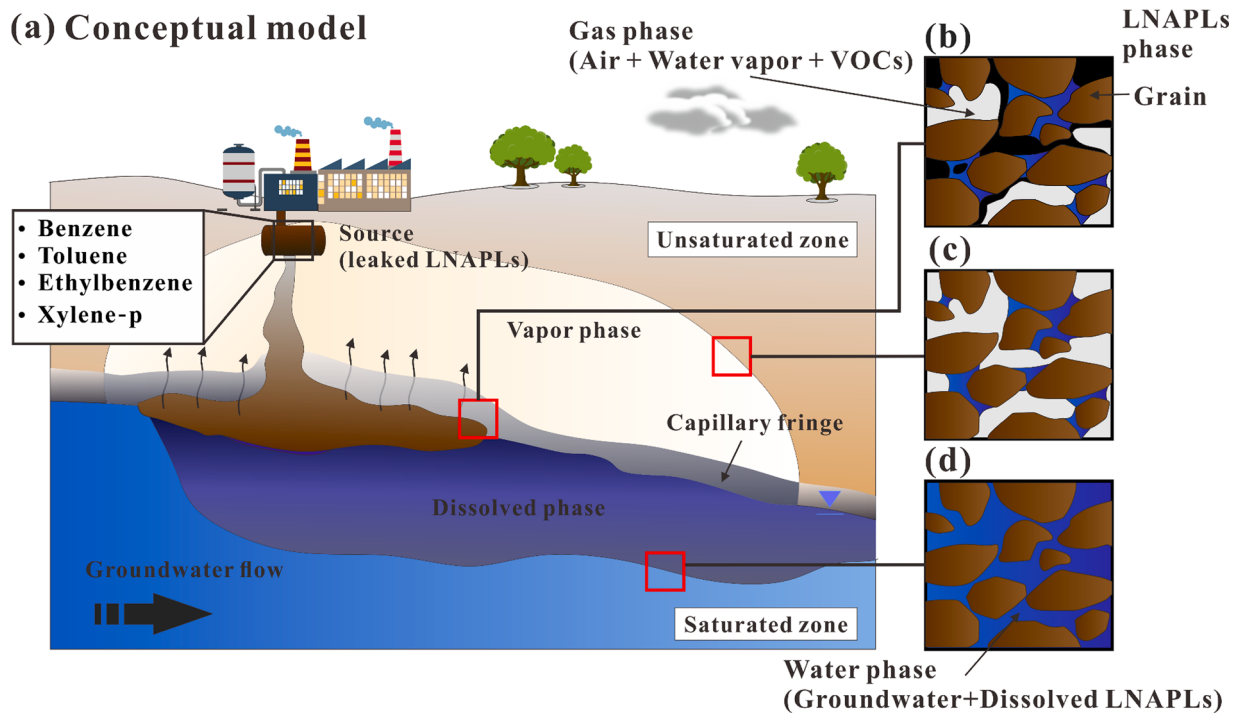


Fig. 1. (a) A conceptual model delineating LNAPLs leakage and transport through both unsaturated and saturated zones. (b) The capillary fringe zone where gas phase (air and VOCs), water phase (residually trapped water), and LNAPLs phase coexisted. (c) In the unsaturated zone, the gas phase primarily existed while both residually trapped water and vaporized LNAPLs are trapped within pores. (d) In the saturated zone, groundwater exists only with dissolved LNAPLs.

migration in the subsurface (Kim and Corapcioglu, 2003). In this study, three different fluids were accounted for, including gas phase (air, water vapor, and VOCs), LNAPL phase, and aqueous phase (residually trapped water in soil and water dissolved in LNAPLs and air). Because air is a dominant component in the unsaturated zone, the gas phase in the unsaturated zone is referred to the air phase. In addition, the aqueous phase in the saturated zone is referred to the water phase.

2.1. Model domain, initial, boundary conditions, and sources

A base-case LNAPLs spill and remediation scenario which can be realized with four modeling stages was considered. The 1st stage is for the development of a sloping groundwater flow, the 2nd stage is the spill of LNAPLs, 3rd stage is the migration of LNAPLs under the ambient condition, and the 4th stage is the active remediation of the LNAPLs.

2.1.1. Groundwater flow (1st stage model)

The model domain shown in Fig. 2a represents the dimensions of 100 m by 50 m, consisting of horizontally discretized 66 grid-blocks per each of 50 layers ($\Delta z=1$ m) from the surface to the bedrock. Porosity and permeability (k_x) were uniformly assigned to 0.3 and $4.0 \times 10^{-13} \text{ m}^2$ to the model domain, respectively; the anisotropy (k_z/k_x) was assigned to 0.25, assuming horizontal arrangement of soil deposits in the sedimentary layers (Nielsen et al., 1986; Scholes et al., 2007).

The groundwater flow model was developed in two stages. In the 1st stage model, a sloping groundwater table was developed, which was the initial condition of the 2nd stage model simulating LNAPL migration. In the 1st stage model, the constant atmospheric temperature and pressure (25°C and $1.01 \times 10^5 \text{ Pa}$) conditions were assigned to the top boundary (Fig. 2a). Immediately below the top boundary grid-blocks, an infiltration rate of $4.64 \times 10^{-6} \text{ kg/s}$ was assigned to represent infiltrating precipitation; the assigned infiltration rate was estimated from the annual average precipitation in the Republic of Korea. Both lateral boundaries were assigned to the constant pressure boundary with pressure gradients (left boundary: $1.01 \times 10^5 \text{ Pa}$ (top) to $3.72 \times 10^5 \text{ Pa}$ (bottom) and right boundary: $1.01 \times 10^5 \text{ Pa}$ (top) to $3.52 \times 10^5 \text{ Pa}$ (bottom)).

The pressure difference between left and right boundaries is attributed to the groundwater flow. The bottom boundary was set to no-flow owing to the presence of low-permeability base rock. Initially, the flat groundwater table was at a depth of 21 m by setting water saturation of 0.2 and 0.99 at unsaturated and saturated zones, respectively. After simulating the 1st stage model, the equilibrated sloping groundwater table was achieved, and the resulting average groundwater velocity was 0.022 m/day.

2.1.2. LNAPLs spilling and remediation (2nd, 3rd, and 4th stage models)

In the 2nd stage model (LNAPL-spilling stage), the LNAPLs composed of BTEX spilled for 5 years from the point source, located 37.5 m away from the left boundary and 4.5 m below the ground surface (Fig. 2b). The spilling constantly occurred at a rate of $1.65 \times 10^{-5} \text{ kg/s}$ per each BTEX component. The primary direction of the LNAPL plume migration is the gravitational direction at this stage, while its migration is governed by the relative saturation of air and LNAPLs in the unsaturated zone. At the end of the 2nd stage, the LNAPL-spilling ceased. During the 3rd stage (natural redistribution stage) that lasted for 2 years, the spilled LNAPLs were naturally redistributed in the capillary fringe, primarily spreading above the sloping groundwater table (Fig. 2c). Finally, in the 4th stage (remediation stage) that lasted for 0.5 years, both the steam injection (SI) and the multi-phase extraction (MPE) wells were installed to remediate the LNAPLs (Fig. 2d). The SI well was installed below the LNAPL source at a depth between 20 m and 22.5 m, where 100°C steam was injected at a rate of $6 \times 10^{-4} \text{ kg/s}$ to aid phase transfer from high-viscosity liquid LNAPLs to low-viscosity vapor LNAPLs. In addition, the MPE well was installed horizontally 5 m away from the LNAPL source at a depth between 16 m and 20 m to extract both vapor and liquid phases of LNAPLs. At the MPE well, a constant bottom-hole pressure (BHP) of $8 \times 10^4 \text{ Pa}$ was assigned, which was approximately equivalent to the rate of $-5.7 \times 10^{-3} \text{ kg/s}$. Finally, four modeling stages representing groundwater flow, LNAPL-spilling, natural redistribution, and remediation were simulated using TOUGH2-TMVOC (Pruess and Battistelli, 2002).

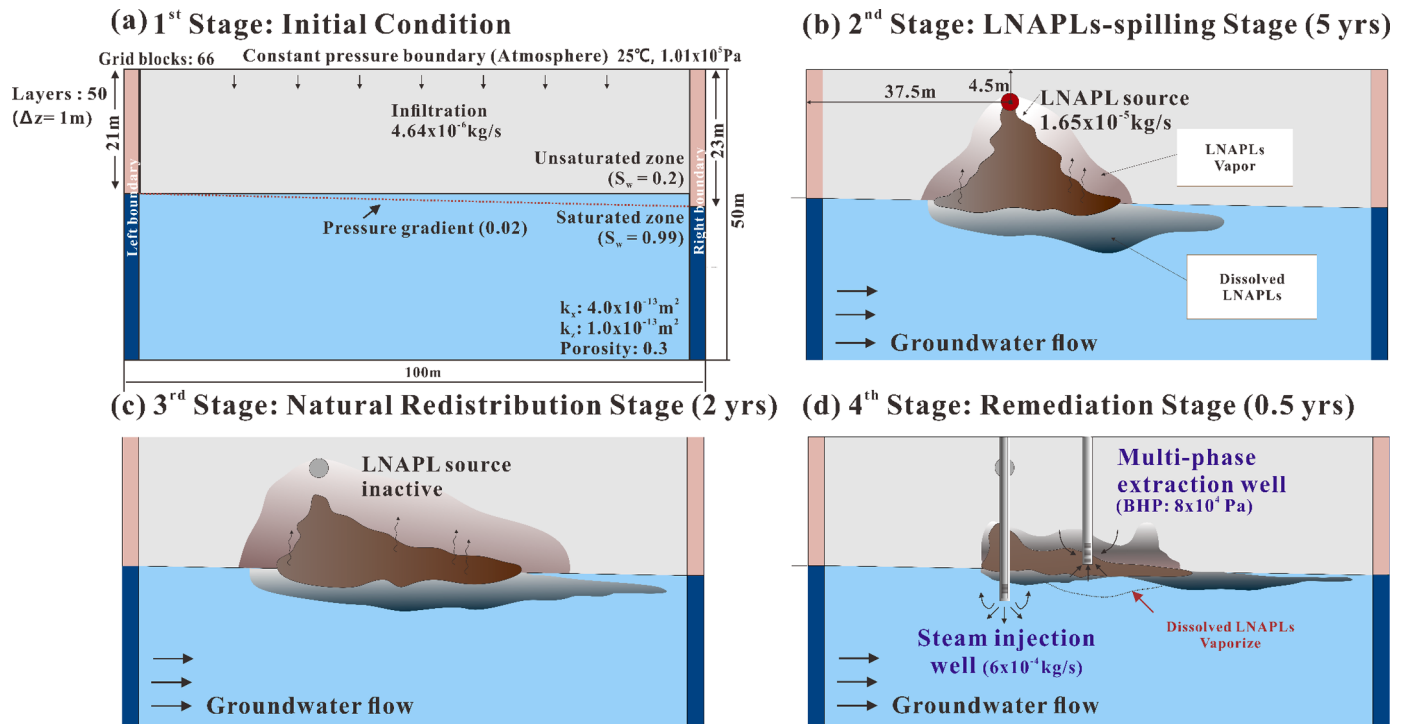


Fig. 2. (a) 1st stage: the initial condition simulating for sloped groundwater table and infiltration, (b) 2nd stage: LNAPLs-spilling stage, (c) 3rd stage: natural distribution, and (d) 4th stage: remediation stage.

2.2. Parameters for 3-Phase flow systems and thermodynamic properties of LNAPLs

Two of the major aspects governing the dynamic behaviors of LNAPLs are capillary pressure and relative permeability; these are parameters characterizing multi-phase fluid systems in porous media. Such multi-phase effects spatiotemporally vary depending on the saturation of the individual fluid phase and the characteristics of porous media (Lenhard and Parker, 1987; Miller et al., 1998).

The capillary pressure caused by the relative difference in wettability and interfacial tension between two immiscible fluids is one of the major factors governing LNAPL migration in unsaturated zones (Anderson, 1987; Hassanizadeh and Gray, 1993). LNAPLs are the intermediate wetting phases among the three phases (air, water, and LNAPLs). Therefore, when the LNAPLs come into contact with air, the LNAPLs and air are treated as wetting and non-wetting fluids, respectively. Whereas, at the LNAPLs contacting with water, the LNAPLs and water are treated as non-wetting and wetting fluids, respectively. In the present study, the 3-phase capillary pressures were calculated from Parker's 3-phase function, which was derived to solve multi-phase organic contaminant transport in the unsaturated zone (Kaluvarachchi and Parker, 1990; Wu and Forsyth, 2001). The capillary pressures were calculated at the interfaces between the air-LNAPLs ($P_{c,AL}$), LNAPLs-water ($P_{c,LW}$), and the water-air ($P_{c,WA}$). The mathematical function for 3-phase capillary pressures are presented in Table 1, where ρ_w is the water density dependent on subsurface pressure and temperature (Pruess and Battistelli, 2002), and g is the gravitational acceleration (9.81 m/s^2). The α_{AL} , α_{LW} , and n are empirically determined parameters, and S_m is the residual wetting phase (water) saturation. Finally, S_w and S_L represent the saturation of water and LNAPLs, respectively.

While groundwater, LNAPLs, and air concurrently migrate through porous media, their movements are restricted according to variability in relative permeability; consequently, the average velocities of the three competing fluids decrease (Alizadeh and Piri, 2014; Bradford et al., 1997). In this study, the 3-phase relative permeabilities (k_r) were calculated from the Stone I model, which is commonly used for water-wet porous media (Baker, 1988; Blunt, 2000; Stone, 1970). In the Stone I model, the relative permeabilities of air (k_r^A), water (k_r^W) and LNAPLs (k_r^L) are functions of the relative saturation of each fluid phase, as shown in Table 1. Here, S_A , S_w , and S_L are the saturation of air, water, and LNAPLs, respectively. In addition, S_{Ar} , S_{Wr} , and S_{Lr} representing residual saturations of air, water, and LNAPLs, respectively, were

obtained from Erning et al. (2009), Juanes (2003), and Pruess (2003), respectively. Finally, n is an adjustable fitting parameter.

In addition to both capillary pressure and relative permeability, accurate prediction of LNAPL migration in the subsurface requires the computation of thermophysical properties in individual BTEX component and their mixtures. Hence, phase transfers (volatilization, condensation, and dissolution of BTEX), which are dependent on vapor pressure, water solubility, and Henry's constant, were accounted for at the designated temperature, pressure, and chemical compositions (Table 2). Corresponding computations are referred to Pruess and Battistelli (2002), and Reid et al. (1987).

3. Workflow

The workflow followed the 4 steps: (1) sampling, (2) development of full-physics numerical model-based surrogate model, (3) global sensitivity analysis, and (4) Monte Carlo prediction.

3.1. Sampling and development of surrogate polynomial chaos expansion model

Latin Hypercube (LH) sampling is a common *space-filling* method that enables uniform sampling with reasonable computational costs (Viana, 2013). In multi-dimensions, the sampling method searches the sampling points (p) within orthogonal grids where each non-overlapped set of rows and columns satisfied the space-filling property. The sampled combination of n number of input factors ($\mathbf{x} = [x_1, x_2, \dots, x_n]$) and corresponding response (Y) computed from the full-physics numerical

Table 2
Calculated Thermodynamic Properties of BTEX Components.

Properties	Benzene	Toluene	Ethylbenzene	Xylene-p
Density (g/cm^3 , 20 °C)	0.88	0.87	0.87	0.86
Molecular weight (g/mol)	78.114	92.141	106.168	106.168
Boiling Point (°C)	80.0	110.6	136.1	138.3
Saturated Vapor Pressure (Pa, 25 °C)	12,523.50	3788.51	1277.48	1169.47
Water Solubility (g/L)	0.411	0.101	0.026	0.030
Henry Constant ($\text{atmm}^3/\text{mole}$)	5.55×10^{-3}	6.64×10^{-3}	7.88×10^{-3}	5.18×10^{-3}
Viscosity (cP, 25 °C)	0.61	0.55	0.64	0.61

Table 1
Two Empirical Functions for Multi-phase Flow System.

Model	Functions and constant parameters
Capillary pressure (Lenhard and Parker, 1987)	Functions $P_{c,AL} = -\frac{\rho_w g}{\alpha_{AL}} \left[(\bar{S}_{Liq})^{-\frac{1}{m}} - 1 \right] \frac{1}{n}$ $P_{c,LW} = -\frac{\rho_w g}{\alpha_{LW}} \left[(\bar{S}_w)^{-\frac{1}{m}} - 1 \right] \frac{1}{n}$ $P_{c,WA} = P_{c,LW} - P_{c,AL}$ $\bar{S}_{Liq} = \frac{S_w + S_L - S_m}{1 - S_m}$ $\bar{S}_w = \frac{S_w - S_m}{1 - S_m}, m = 1 - \frac{1}{n}$ Parameters $n = 1.84 \quad S_m = 0.2 \quad \alpha_{LW} = 10 \quad \alpha_{AL} = 11$
Relative permeability (Stone, 1970)	Functions $k_r^A = \left[\frac{S_A - S_{Ar}}{1 - S_{Wr}} \right]^n$ $k_r^W = \left[\frac{S_w - S_{Wr}}{1 - S_{Wr}} \right]^n$ $k_r^L = \left[\frac{1 - S_A - S_w - S_{Lr}}{1 - S_A - S_{Wr} - S_{Lr}} \right] \left[\frac{1 - S_{Wr} - S_{Lr}}{1 - S_{Wr} - S_{Lr}} \right] \left[\frac{(1 - S_A - S_{Wr} - S_{Lr})(1 - S_w)}{(1 - S_{Wr})} \right]^n$ Parameters $n = 3.0 \quad S_{Wr} = 0.2 \quad S_{Ar} = 0.01 \quad S_{Lr} = 0.05$

LNAPL model composes the training dataset used to build a surrogate model. From $\mathbf{x}$ and Y , the surrogate polynomial chaos expansion (PCE) model can be established following Blatman and Sudret (2008):

$$Y = M(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}^n} \lambda_{\alpha} \Psi_{\alpha}(\mathbf{x}) \quad (1)$$

where λ_{α} is the deterministic chaos expansion coefficient, and $\Psi_{\alpha}(\mathbf{x})$ is the multivariate orthonormal basis polynomial. Assuming the independence between the input factors in $\mathbf{x}$, the $\Psi_{\alpha}(\mathbf{x})$ is established by the multiplication of univariate polynomials of each input factor (x_i),

$$\Psi_{\alpha}(\mathbf{x}) = \prod_{i=1}^n \psi_{\alpha_i}(x_i) \quad (2)$$

where each ψ_{α_i} is the univariate polynomial of degree α_i with respect to the i^{th} input factor. The type of the ψ_{α_i} is selected based on the type of probability density function (PDF) of x_i . Considering the uniform and Gaussian PDFs, the $\psi_{\alpha_i}(x_i)$ are selected to normalized Legendre and Hermitian polynomials, respectively (Xiu, 2010). Because the input factors employed in this study were all assumed to be uniform distribution, the type of the univariate polynomial ($\psi_{\alpha_i}(x_i)$) was regarded as the Legendre polynomial.

For the practical computation of the targeted PCE from the Legendre polynomials of $\mathbf{x}$, which was linearly transformed into $[-1, 1]^n$, $M(\mathbf{x})$ would be truncated up to the total degree of the $\Psi_{\alpha}(\mathbf{x})$, ω :

$$M(\mathbf{x}) \cong M_T(\mathbf{x}) = \sum_{\alpha \in A^{n,\omega}} \lambda_{\alpha} \Psi_{\alpha}(\mathbf{x}), \quad \mathbf{x} \sim u([-1, 1]^n) \quad (3)$$

where the truncated PCE, $M_T(\mathbf{x})$, is then remained as the finite terms corresponding to subsets of $A^{n,\omega}$, which refers to the truncation set, satisfying $|\alpha| = \sum_{i=1}^n \alpha_i \leq \omega$. To compute the designated number of Legendre polynomials ($\Psi_{\alpha}(\mathbf{x})$) and the corresponding chaos expansion coefficients (λ_{α}) in $A^{n,\omega}$, the cardinality of $A^{n,\omega}$ should be determined to be $P = (n + \omega)!/n!\omega!$. Here, n and ω represent the number of input factors and the total degree of the $\Psi_{\alpha}(\mathbf{x})$, respectively. After determining multivariate basis polynomials ($\Psi_{\alpha}(\mathbf{x})$), λ_{α} can be calculated through either intrusive or non-intrusive schemes (Song et al., 2019). In the present study, least-squares minimization (LSM) in the non-intrusive scheme is utilized to calculate λ_{α} . Here, to achieve the λ_{α} accurately, the number of sampling points (p) from the LH sampling should be greater than P (Blatman and Sudret, 2008, 2010b).

3.2. Global sensitivity analysis

The Sobol global sensitivity analysis is a type of variance-based sensitivity analysis, which is applicable when the input variables are independent. The Sobol scheme practically results in quantified sensitivity indicators, Sobol indices, from the Monte Carlo simulation. Blatman and Sudret (2010a) and Sudret (2008) provided a method for the straightforward computation of Sobol indices from the PCE. This approach significantly reduces the computational cost to calculate the Sobol indices, requiring 2 to 3 orders of magnitude fewer simulations than the Monte Carlo simulation.

For the PCE of the multivariate Legendre polynomials, Eq. (3) can be expressed as:

$$Y = f(X) \cong M_T(X) = \sum_{j=0}^{P-1} \lambda_j \Psi_j(X), \quad X \sim u([-1, 1]^n). \quad (4)$$

The mean and variance of the PCE were computed based on the orthogonality of the basis polynomial as follows:

$$\bar{Y} = E[f(X)] = \lambda_0 \quad (5)$$

$$V_{PCE} = \text{Var} \left[\sum_{j=0}^{P-1} \lambda_j \Psi_j(X) \right] = \sum_{j=0}^{P-1} \lambda_j^2 E[\Psi_j^2(X)] \quad (6)$$

Eq. (4) can be expanded by Sobol decomposition after derivation (Sudret, 2008), as shown below:

$$\begin{aligned} f(X) &\cong M_T(X) \\ &= \lambda_0 \\ &+ \sum_{i=1}^n \sum_{\alpha \in h_1} \lambda_{\alpha} \Psi_{\alpha}(x_i) \\ &+ \sum_{1 \leq i_1 < i_2 \leq n} \sum_{\alpha \in h_{i_1, i_2}} \lambda_{\alpha} \Psi_{\alpha}(x_{i_1}, x_{i_2}) \\ &+ \dots \sum_{1 \leq i_1 < \dots < i_s \leq n} \sum_{\alpha \in h_{i_1, \dots, i_s}} \lambda_{\alpha} \Psi_{\alpha}(x_{i_1}, \dots, x_{i_s}) + \dots \\ &+ \sum_{\alpha \in h_{i_1, i_2, \dots, n}} \lambda_{\alpha} \Psi_{\alpha}(x_1, \dots, x_n). \end{aligned} \quad (7)$$

where each summand $\sum_{\alpha \in h_{i_1, \dots, i_s}} \lambda_{\alpha} \Psi_{\alpha}(x_{i_1}, \dots, x_{i_s})$ are orthogonal polynomials produced by $x_{i_1}, \dots, x_{i_s}$, and $h_{i_1, \dots, i_s}$ is the set in α , which indicates that only the $x_{i_1}, \dots, x_{i_s}$ are involved in the summand. As a result of squaring both the left and right sides of Eq. (7), the Sobol indices combined with the PCE ($SI_{i_1, \dots, i_s}$) for any $h_{i_1, \dots, i_s}$ can be obtained as the proportion of the partial variance to the total variance:

$$SI_{i_1, \dots, i_s} = \sum_{\alpha \in h_{i_1, \dots, i_s}} \lambda_{\alpha}^2 E[\Psi_{\alpha}^2] / V_{PCE} \quad (8)$$

Typically, two types of Sobol indices are used: the Sobol 1st-order indices $\left(SI_{i_1}^{1^{\text{st}}} = \sum_{\alpha \in h_{i_1}} \lambda_{\alpha}^2 E[\Psi_{\alpha}^2] / V_{PCE} \right)$ capable of showing the only influence of each independent variable on output and the Sobol total-order indices $\left(SI_{i_1}^T = \sum_{i_1 \in \{i_1, \dots, i_s\}} SI_{i_1, \dots, i_s} \right)$ considering additional effects by the interactions of the independent variables.

4. Results

4.1. Full-physics numerical LNAPLs transport model: base-case

4.1.1. Spatiotemporal distribution of air, LNAPLs, and water

Simulation results for phase saturation (S_L , S_A , and S_W) at the end of 2nd ($t=5$ years), 3rd ($t=7$ years), and 4th ($t=7.5$ years) stages are shown in Fig. 3. The sources for spilling LNAPLs are marked as a red circle, and the white line represents the groundwater table.

During the 2nd stage (LNAPL-spilling stage), LNAPLs migrated downward through the unsaturated zone until they reached the capillary fringe above the groundwater table (Fig. 3a). The LNAPL plume in the unsaturated zone expanded symmetrically, owing to the homogeneous permeability and dispersion; the average S_L in the unsaturated zone was 0.11. The corresponding S_A and S_W were concurrently changed to 0.65 and 0.24, respectively (Fig. 3b and 3c). Since S_A and S_W were 0.75 and 0.25 before LNAPLs spill, changes in S_A during the 2nd stage indicated that the spilled LNAPLs primarily displaced the low-viscosity air in the pore space. Immediately above the groundwater table, the LNAPLs accumulated in the capillary fringe; the capillary fringe served as a primary flow path that allowed the LNAPLs to spread horizontally (Mayer, 2005). In the model, before the LNAPLs spill, the thickness of the capillary fringe was approximately 3 m with associated S_A and S_W of 0.45 and 0.55, respectively (Fig. 3b and 3c). Once the LNAPL plume reached the capillary fringe, it primarily displaced low-viscosity air; the S_L increased from 0 to 0.45 (Fig. 3a) while S_A decreased from 0.45 to almost 0.0 (Fig. 3b). Additionally, due to the difference in both density and wettability between LNAPLs and groundwater, the downward movement of LNAPLs below the groundwater table was hindered. Instead, the LNAPLs migrated approximately 9 m along the sloping

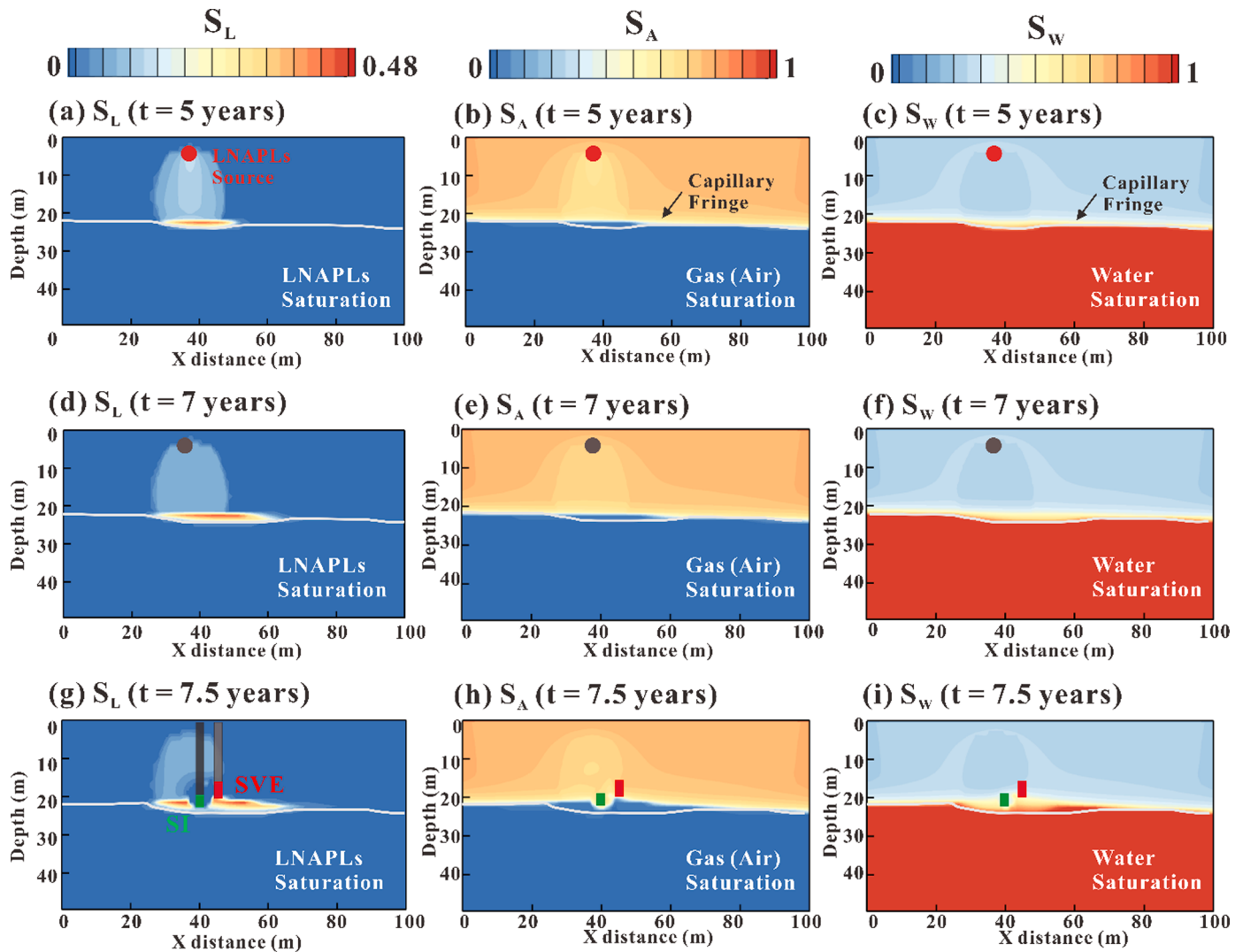


Fig. 3. (a) S_L , (b) S_A , and (c) S_W at the end of the 2nd stage (5 years). (d) S_L , (e) S_A , and (f) S_W at the end of the 3rd stage (7 years). (g) S_L , (h) S_A , and (i) S_W at the end of the 4th stage (7.5 years). The white lines indicate the groundwater table.

groundwater table. Finally, owing to the counter buoyancy force acting on the LNAPL plume, the groundwater table was suppressed approximately 1.5 m, and consequently, the capillary fringe was thickened.

During the 3rd stage (natural redistribution stage), S_L in the unsaturated zone uniformly dropped to 0.052 close to the assigned S_{Lr} (=0.05) (Table 1). The corresponding k_r^L calculated from the Stone I model was almost 0, indicating that most LNAPLs remained as the residually trapped phase in the unsaturated zone (Fig. 3d). Above the groundwater table, the mobile LNAPLs continuously accumulated, while the LNAPL plume spread horizontally above the sloping groundwater table. At the end of the 3rd stage, the maximum S_L of the floating LNAPL plume increased to 0.462. In addition, S_A in the unsaturated zone nearly returned to the initial saturation because the air re-filled pore spaces where movable LNAPLs were left (Fig. 3e).

Finally, the S_L , S_A , and S_W are depicted at the end of the 4th stage (remediation stage) where the remediation scheme was implemented (Fig. 3g, 3h, and 3i). As mentioned in Section 2.1.2, two remediation wells including steam injection (SI) and multi-phase extraction (MPE) wells, were operated; the screen depths of each well are highlighted in red and blue boxes, respectively. At the end of the remediation stage, the S_L of the floating LNAPL plume rapidly decreased to 0.22, mainly due to extraction of the movable LNAPLs and phase transfer from liquid LNAPLs to volatilized LNAPLs by high-temperature steam injection

(Fig. 3g). The pressure gradient induced by the MPE well aided to extract low-viscosity volatilized LNAPLs. Additionally, the pressure gradient induced by the two remediation wells distorted the capillary fringe. For instance, adjacent to the SI well, a hole, where S_A was increased to 0.69 due to both volatilized LNAPLs and injected steam, was developed (Fig. 3h). In the unsaturated zone, the simultaneous operation of the two remediation wells was able to remove trapped LNAPLs; S_L was decreased to below 0.05 (Fig. 3g). At the bottom of the MPE well, both S_L and S_W increased, indicating that both LNAPLs and groundwater were effectively extracted by the MPE well (Fig. 3g and 3i).

4.1.2. BTEX in air, LNAPLs, and water phases

The distributions of BTEX (benzene, toluene, ethylbenzene, and xylene-p) at the end of the 3rd stage were delineated to elucidate the different behaviors of the spilled BTEX in the subsurface, based on their thermophysical properties (Fig. 4). X_L , X_A , and X_W represent the mole fraction of BTEX existing in LNAPLs, air, and water, respectively.

In the LNAPL phase, the BTEX components were evenly distributed, although their average X_L values were different due to the difference in molar mass (Fig. 4a). The X_L of BTEX were 0.3, 0.25, 0.22, and 0.22, respectively. The mole fractions of BTEX in the air phase are delineated in Fig. 4b, which reveals a clear difference depending on the volatilization of BTEX. The degree of BTEX volatilization was related to the

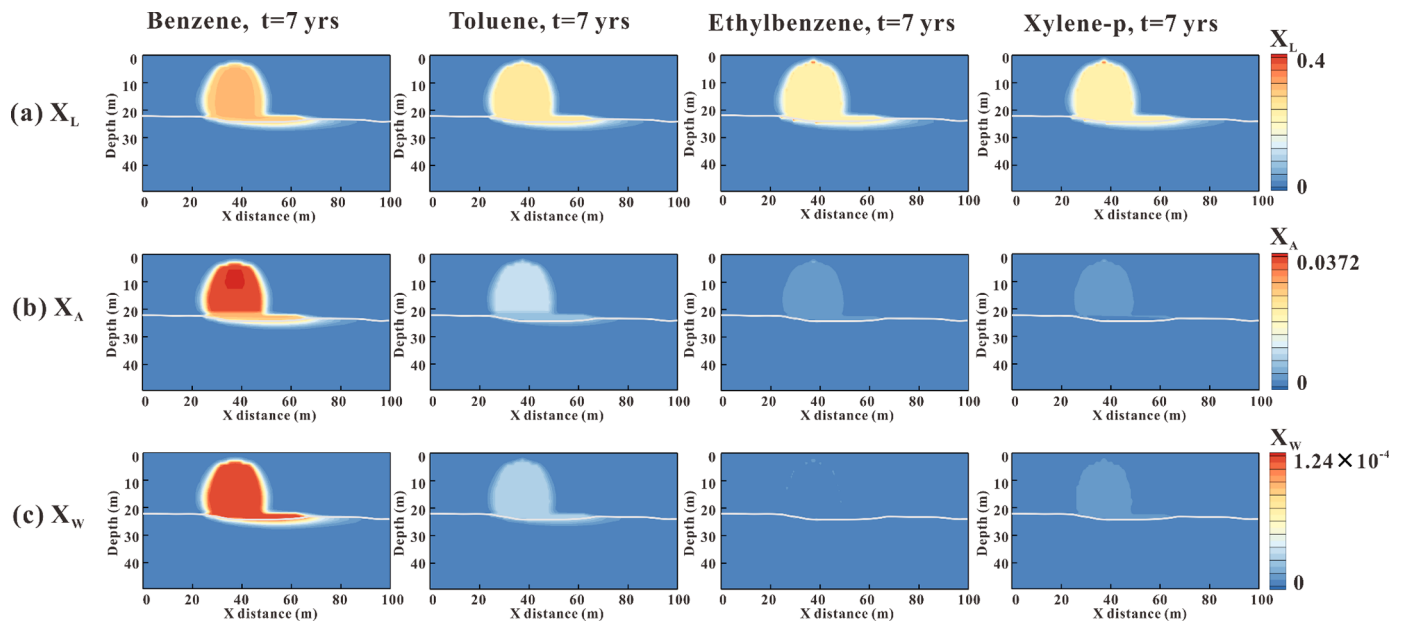


Fig. 4. Distribution of BTEX mole fractions existing in (a) LNAPLs phase (X_L), (b) air phase (X_A), and (c) water phase (X_W) at the end of the 3rd stage (7 years).

saturated vapor pressure (Table 2). Benzene characterizing the largest saturated vapor pressure (12,523.50 Pa) showed the maximum X_A (0.0372) which was 8.5 times greater than that of xylene-p (0.00436) having the lowest saturated vapor pressure (1169.47 Pa). Additionally, benzene, which had the largest solubility (0.411 g/L) among the BTEX, showed the maximum X_W of 1.24×10^{-4} (Fig. 4c). In summary, among BTEX, Benzene was the dominant contaminant due to its high saturated vapor pressure and water solubility.

4.1.3. Removal efficiency of LNAPLs

The BTEX distribution at the end of the 4th stage (remediation stage) after 7.5 years is shown in Fig. 5. At the end of the remediation stage, the X_L of benzene was the lowest (0.301) among the BTEX components, indicating that the largest amount of removed components was benzene (Fig. 5a). The largest removal of benzene was due to effective phase transfer of benzene from liquid LNAPLs to volatilized LNAPLs,

responding to steam injection by the SI well. The temperature of the injected steam was approximately 100 °C and the boiling temperature of benzene was 80 °C, whereas those for other components were over 100 °C (Table 2). The results of such active volatilization in benzene were shown as the largest X_A among the BTEX components (Fig. 5b); the maximum X_A of benzene after the remediation was 0.31, while the average X_A of other components remained at low quantities (T: 0.11; E: 0.08 and X: 0.08) (Fig. 5b). The increased X_A of benzene accelerated the benzene transport to the MPE well, by reducing the viscosity and increasing the relative permeability. In contrast to X_A of benzene, X_W of benzene was not significantly reduced by the remediation (Fig. 5c). This was due to the relatively small viscosity and small relative permeability of water. At the unsaturated zone, water was residually trapped in the pores, and BTEX components dissolved in such residual water could not be extracted by the MPE well.

To evaluate the removal efficiency (RE), a mass of component Z in

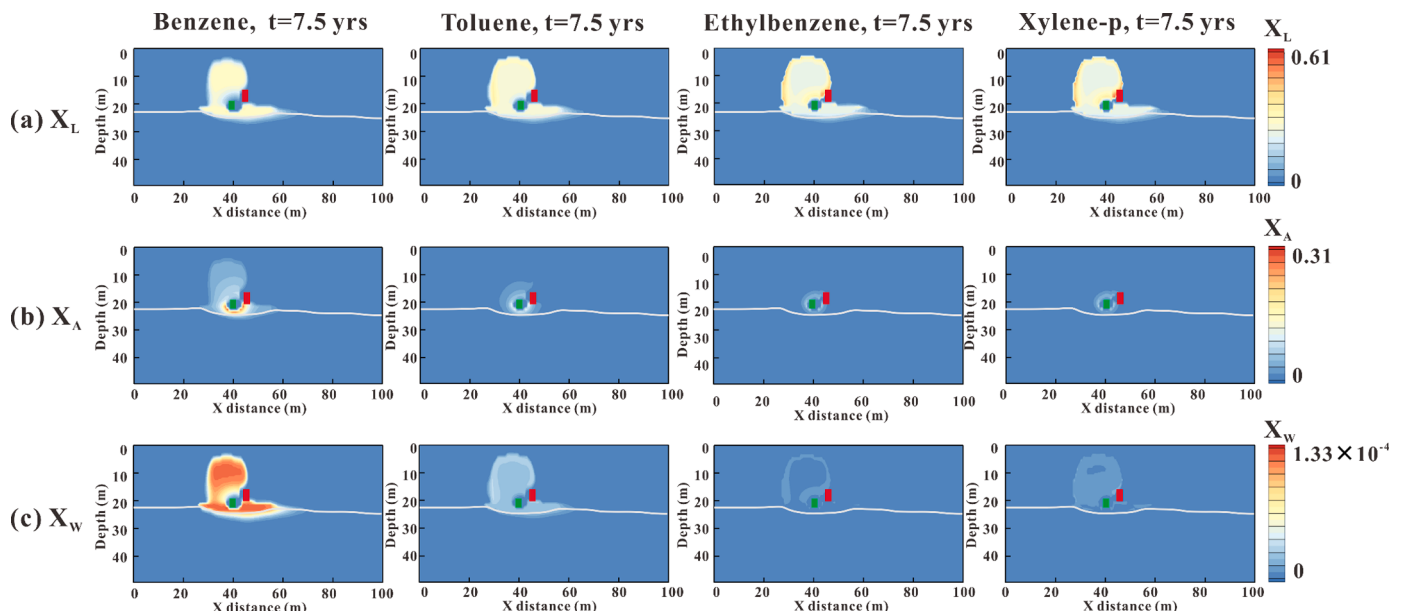


Fig. 5. Distribution of BTEX mole fractions existing in (a) LNAPLs phase (X_L), (b) air phase (X_A), and (c) water phase (X_W) at the end of the 4th stage (7.5 years).

phase P (Z_P^M) before and after remediation (at the end of the 3rd and 4th stage) was calculated. The RE ($Z_P^{RE} = 1 - Z_P^{M^{4th}} / Z_P^{M^{3rd}}$) was then defined, and the REs of individual components (B, T, E, and X) in LNAPLs, air, and water phases were calculated (bar graph in Fig. 6). Additionally, the REs ($Z_P^{RE} = 1 - Z_P^{M^{4th}} / Z_P^{M^{3rd}}$) of individual components (B, T, E, and X) in the total (LNAPLs+air+water) phase were delineated (red circle in Fig. 6).

The REs of BTEX components between the LNAPL (B_L^{RE} , T_L^{RE} , E_L^{RE} , and X_L^{RE}) and total (B_T^{RE} , T_T^{RE} , E_T^{RE} , and X_T^{RE}) phases were nearly the same, with values of approximately 0.23, 0.13, 0.06, and 0.06, respectively (Fig. 6a). This indicated that the BTEX primarily existed as the LNAPL phase, although phase transfer to air and water had occurred. The removed masses of BTEX in the LNAPL phase are proportional to their volatilities (Table 2); here, benzene showed the largest volatility. As more benzene changed to the air phase, the mobility of benzene in the air phase was enhanced. Then, the MPE well accelerated the removal rate of benzene in the air phase. Consequently, the B_A^{RE} (0.4) exceeded B_T^{RE} (0.23). Nevertheless, T_A^{RE} (−0.08), E_A^{RE} (−0.84), and X_A^{RE} (−0.97) showed negative values, indicating that lower volatile components were not efficiently removed by the MPE well after they had been volatilized. (Fig. 6b). The removal efficiencies of benzene, toluene, ethylbenzene, and xylene-p in the water phase (B_W^{RE} , T_W^{RE} , E_W^{RE} , and X_W^{RE}) showed the smallest deviation from 0 (Fig. 6c). Benzene and toluene with high solubility specifically showed positive B_W^{RE} (0.14) and T_W^{RE} (0.01), respectively but ethylbenzene and xylene-p with low solubility showed negative E_W^{RE} (−0.14) and X_W^{RE} (−0.17), respectively. In summary, the components showing effective phase change represented high removal efficiency.

4.2. Surrogate polynomial chaos expansion LNAPLs models

4.2.1. 4 cases: permeability and the location of steam injection wells

Including base-case (Case I), three additional cases were designed to evaluate the effect of permeability (k) and the location of the steam injection (SI) well (Table 3). In Case II, $k_x (=4.0 \times 10^{-13} \text{ m}^2)$ and $k_z (=1.0 \times 10^{-13} \text{ m}^2)$ was equal to Case I, but the location of the SI well was shifted to the right of the MPE well (Fig. 7a and 7b). At both Case I and II represented by high k , LNAPL plumes spread widely above the groundwater table while residually trapped LNAPLs remained in the unsaturated zone. The REs of total phase BTEX (B_T^{RE} , T_T^{RE} , E_T^{RE} , and X_T^{RE}) for Case I were similar to those for Case II, indicating that the relative location of SI to the MPE well did not influence a degree of remediation efficiency much (Fig. 7e and 7f); the B_T^{RE} (0.25 and 0.20), T_T^{RE} (0.13 and 0.10), E_T^{RE} (0.04 and 0.04), and X_T^{RE} (0.04 and 0.04) were predicted at Case I and II, respectively.

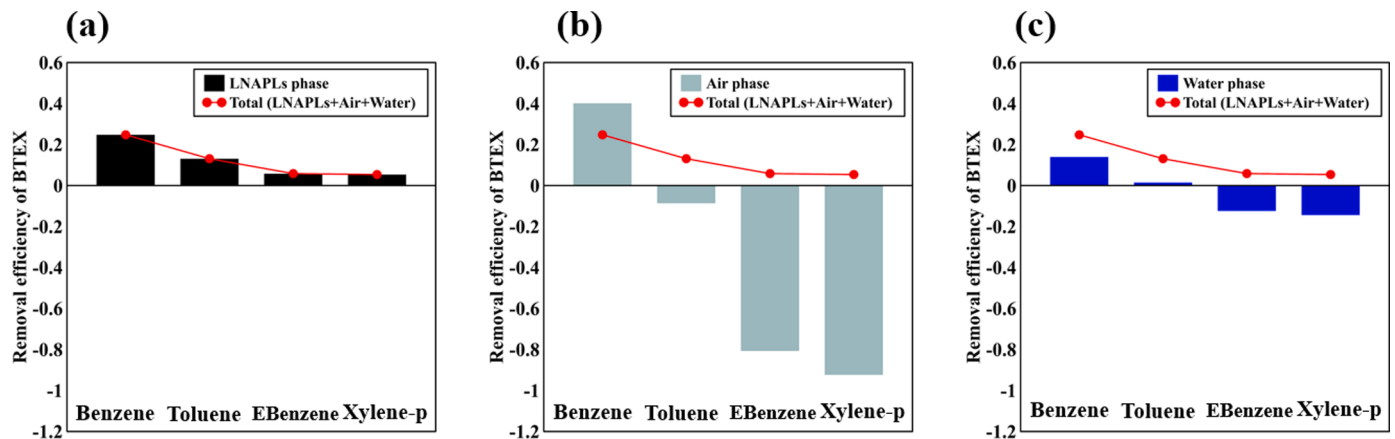


Fig. 6. Bar graphs indicate removal efficiencies of BTEX existing in (a) LNAPLs, (b) air, and (c) water phase, respectively. The red symbols denote the total removal efficiency for BTEX.

Table 3

Four Cases and Six Input Factors.

Case	Permeability (m ²)	Lateral distance from Source to SI well (m)	Input factors	Min	Max
Case I (Base case)	4×10^{-13}	0	x_1 : Screen depth of MPE well (m) x_2 : Screen depth of SI well (m)	16 18	24 27
Case II	4×10^{-13}	10	x_3 : Distance of MPE well from SI well (m)	0	10
Case III	4×10^{-14}	0	x_4 : Bottom-hole pressure of MPE well (Pa) x_5 : Steam injection rate from SI well (kg/s)	7×10^4 2×10^{-4}	9×10^4 1×10^{-3}
Case IV	4×10^{-14}	10	x_6 : Steam temperature (°C)	100	150

For Case III and IV, both k_x and k_z decreased 10 times to $4.0 \times 10^{-14} \text{ m}^2$ and 1.0×10^{-14} while relative locations of the SI well were equal to one for Case I and II, respectively (Fig. 7c and 7d). A decrease in k caused local accumulation of the LNAPL plume with preventing horizontal expansion. The small and concentrated LNAPL plume caused less contact with the surrounding air, and thus, the amount of volatilized LNAPLs decreased. Additionally, decreased mobility of LNAPLs due to low k let more LNAPLs remain in the unsaturated zone and impeded LNAPL migration to the MPE well, causing a decrease in B_T^{RE} , T_T^{RE} (Case III and IV in Fig. 7g and 7h).

4.2.2. Input factors and responses for assessing removal efficiency of LNAPLs

For each of the 4 cases, the 6 input factors governing the effectiveness of the remediation wells were selected (Table 3): the screen depths of both the MPE and SI wells (x_1 and x_2), the lateral distance between the MPE and SI wells (x_3), the BHP of the MPE well (x_4), and the injection rate (x_5), and the temperature of the SI well (x_6). The 6 input factors were acknowledged as influencing factors governing the removal efficiency of LNAPLs (McCray and Falta, 1997; Robin and Gillham, 1987; Rogers and Ong, 2000). These input factors were assumed to have a uniform distribution, with minimum and maximum limits (Table 3). Within these ranges, the sampling points representing the different combinations of input factors were selected through LH sampling. For example, the x_1 and x_2 were selected within the vertical sky-blue and red

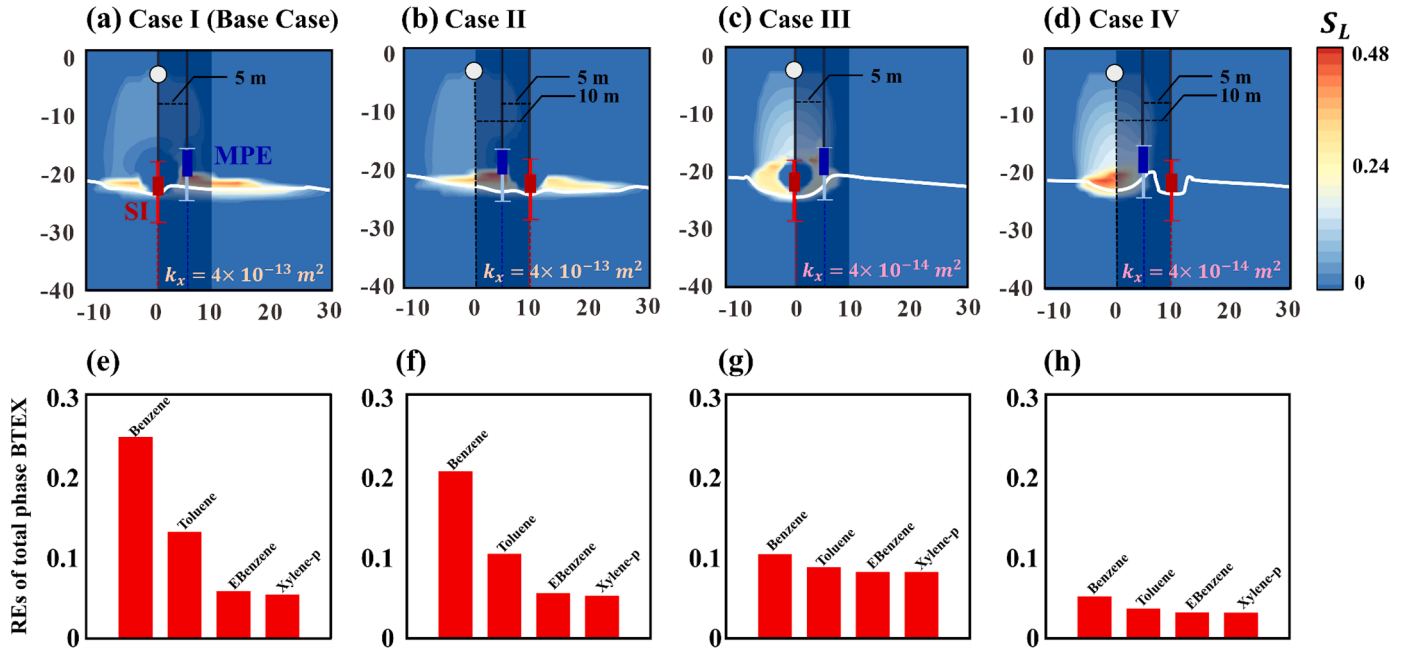


Fig. 7. The S_L distribution of the representative model of (a) Case I, (b) Case II, (c) Case III, and (d) Case IV. Case I and II representing high permeability ($k = 4 \times 10^{-13} \text{ m}^2$) showed the thin and wide spread of the LNAPLs plume above the groundwater table (white line). Case III and Case IV representing low k ($k = 4 \times 10^{-14} \text{ m}^2$) showed concentrated distribution beneath the LNAPLs source (white circle). The shaded area represents the possible horizontal location of the MPE well, determining the x_3 . Corresponding removal efficiency for total phase BTEX was plotted in (e) Case I, (f) Case II, (g) Case III, and (h) Case IV.

scale lines represented in Fig. 7, respectively. Finally, a single sampling point determined a single configuration of wells. In each case, a total of 150 points were sampled (600 points in 4 cases). Based on LH sampling points, a total of 600 full-physics numerical LNAPL transport models were conducted to obtain responses. The responses were the removal efficiencies of the total sum of BTEX ($BTEX_{RE} = 1 - \frac{BTEX_{T_{4th}}}{BTEX_{T_{3rd}}}$) calculated at the end of 4th stage (7.5 years). Then, both input factors and responses served as the training dataset in developing the surrogate polynomial chaos expansion (PCE) models of 4 cases.

4.2.3. Development of surrogate pce model, validation, and test

PCE models predicting $BTEX_{RE}$ were developed using the following procedure. With a given training dataset (e.g. a combination of the 6 input factors and the responses), multi-variate basis functions (Ψ_α) for the PCE models were developed. Preliminary PCE models were then developed by evaluating the determination of the coefficients (λ_α) which were calculated from the least square minimization (LSM) method. However, the preliminary PCE models based on LSM may preserve unnecessary complexities and overfitting problems (Blatman and Sudret, 2010b). To overcome these problems, the preliminary PCE models were improved through the validation process; both an adaptive-sparse algorithm with the least angle regression and a leave-one-out (LOO) cross-validation algorithm were used for the validation.

The adaptive-sparse algorithm is a step-wise procedure that reduces the complexity of the PCE model by eliminating the insignificant Ψ_α among the total P number of Ψ_α (Blatman and Sudret, 2011); the total P ($P = (n + \omega)! / (n! \omega!)$) number of Ψ_α was 210 ($n=6$ and $\omega=4$), 210 ($n=6$ and $\omega=4$), 210 ($n=6$ and $\omega=4$), and 84 ($n=6$ and $\omega=3$) in Case I, II, III, and IV, respectively. After applying the advanced adaptive-sparse method with the least angle regression, the total P number of Ψ_α in preliminary PCE models decreased to 29, 30, 53, and 14 in Case I, II, III, and IV. Once the PCE models were streamlined, the deterministic chaos expansion coefficients (λ_α) were subsequently modified to solve the overfitting problem, by minimizing the LOO error (ϵ_{LOO}) (Blatman, 2009):

$$\epsilon_{LOO} = \frac{\sum_{i=1}^N (M(x^{(i)}) - M^{PCV}(x^{(i)}))^2}{\sum_{i=1}^N (M(x^{(i)}) - \hat{\mu}_y)^2} \quad (9)$$

where N is the number of the training dataset ($N=150$), $M(x^{(i)})$ is the i^{th} response calculated from a sparse PCE model developed by N training data, $M^{PCV}(x^{(i)})$ is the i^{th} response of a sparse PCE model, but the model derived from $N - 1$ training data by excluding the i^{th} training data, and $\hat{\mu}_y$ is the mean of the N responses. Finally, the optimal ϵ_{LOO} values of the sparse PCE models were calculated to be 0.036, 0.044, 0.029, and 0.066 in Case I, II, III, and IV, respectively. Once the optimal ϵ_{LOO} was calculated, the corresponding λ_α values were selected to determine validated PCE models.

In Fig. 8, the $BTEX_{RE}$ calculated from both the validated PCE models and the full-physics numerical LNAPLs transport models are plotted for Cases I, II, III, and IV; here, 150 yellow circles and 50 green triangles indicate the training and test data, respectively. In each case, the validated PCE models developed using 150 training data were tested with 50 test data, which were randomly chosen by the LH sampling. Using 50 test data, the determination coefficient (R^2) and normalized root mean squared error (NRMSE) were calculated to investigate the predictability of the validated PCE models. The predictability of the validated PCE models was the highest in Case IV (Fig. 8d); R^2_{test} and $NRMSE_{\text{test}}$ were 0.930 and 0.067, respectively. Even for Case II showing the lowest predictability (Fig. 8b), the predictability was still acceptable (R^2_{test} : 0.889 and $NRMSE_{\text{test}}$: 0.11), ensuring that the validated PCE models were capable of substituting the responses of the full-physics numerical LNAPLs transport models.

Interestingly, both training and test datasets were distributed differently depending on the magnitude of permeability (k), indicating that even small differences in k significantly influenced $BTEX_{RE}$. At high k values (Cases I and II), both training and test datasets were widely distributed between 0 and 0.55 (Fig. 8a and b), and the widely distributed $BTEX_{RE}$ indicates that the $BTEX_{RE}$ was sensitive to remediation conditions represented by 6 input factors. Dependent on suitable combinations of 6 input factors, the $BTEX_{RE}$ could be maximized. In contrast,

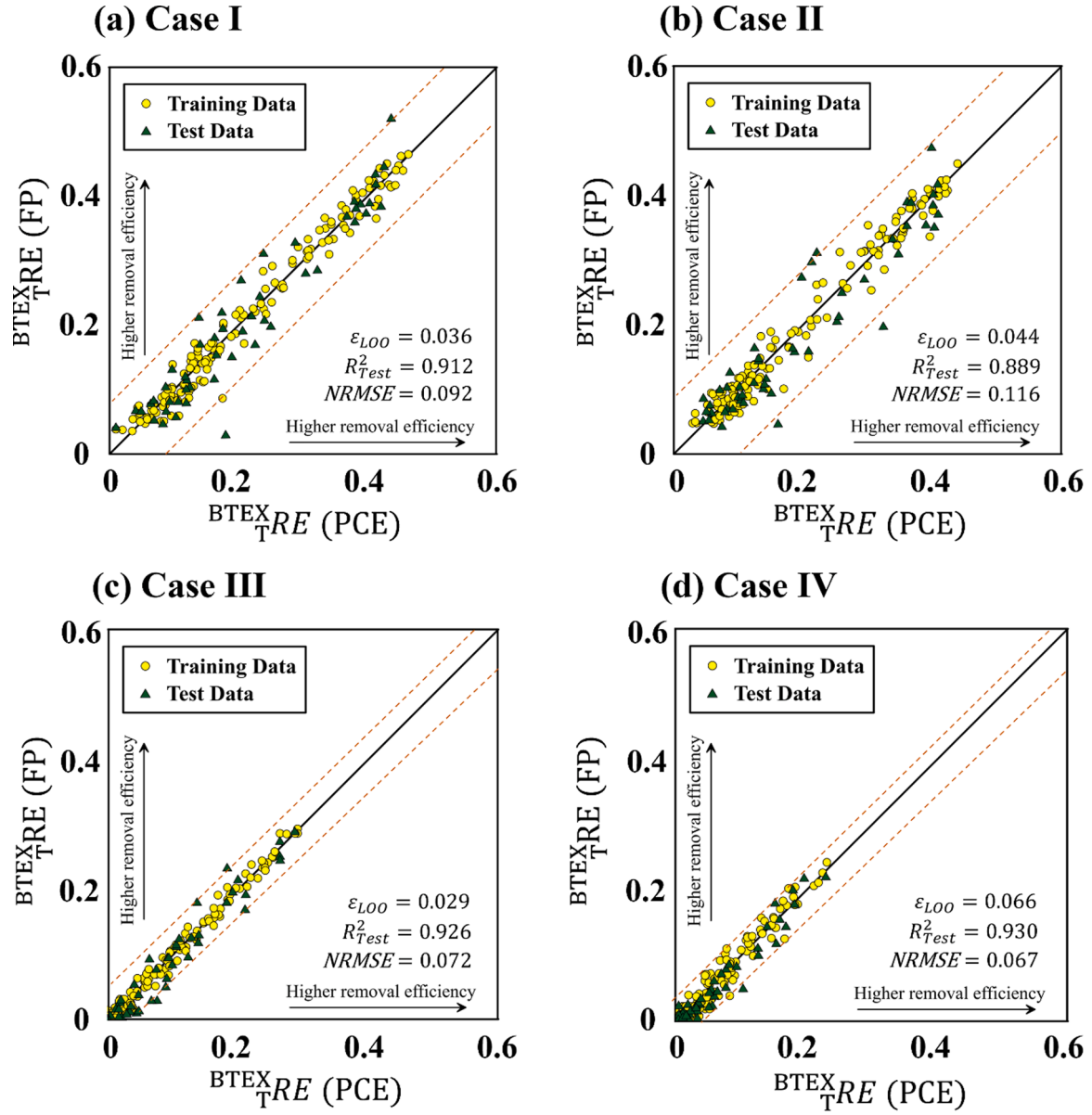


Fig. 8. The removal efficiency calculated from surrogate PCE model (x-axis) and numerical full-physics (FP) numerical model (y-axis) were plotted in 1:1 graph for (a) Case I, (b) Case II, (c) Case III, and (d) Case IV. Orange-dotted lines indicate 95% confidence intervals.

at low k (Cases III and IV), most of the training and test data were leaned to the 0 (small $\frac{BTEX}{T}RE$), denoting that $\frac{BTEX}{T}RE$ cannot be improved significantly by the choice of the 6 input factors. In Fig. 7, the distribution of the LNAPL plume at both high k and low k is shown. At high k , the LNAPL plume spread widely over the groundwater table and even migrated far from both MPE and SI well (Fig. 7a and b). Nevertheless, $\frac{BTEX}{T}RE$ could be maximized dependent on the choice of 6 input factors (Fig. 8a and b). This is because the LNAPLs can be more easily mobilized by remediation wells when k is high. At low k , however, although the LNAPL plume was located close to both the MPE and SI well (Fig. 7c and d), the overall $\frac{BTEX}{T}RE$ was still small because the high pumping rate of the MPE well did not improve the mobility of the LNAPLs (Fig. 8c and d). Such a difference in $\frac{BTEX}{T}RE$ implies that the influence of k on the remediation of the LNAPL plume presumably prevails over any 6 input factors. Therefore, at contamination sites with extremely small k (e.g., clay-dominant or fractured rock sites), changes in 6 input factors may not be able to improve $\frac{BTEX}{T}RE$, significantly.

4.3. Analysis using surrogate pce models

4.3.1. Sobol global sensitivity analysis

Using the validated PCE models, the influence of the input factors (x_1, x_2, x_3, x_4, x_5 , and x_6) on the $\frac{BTEX}{T}RE$ were assessed through Sobol global sensitivity analysis (Case I in Fig. 9a, Case II in Fig. 9b, Case III in Fig. 9c, and Case IV in Fig. 9d). For each case, the Sobol total (T)- and first (1st)-order indices were calculated with 150 realizations of the PCE model, whereas they would be yielded by more than 10^4 realizations with a common Monte Carlo simulation (Sudret, 2008). Sobol T indices (blue bar), which include the effect of intercorrelation between the input factors, are always greater than Sobol 1st indices (purple bar).

Similar to Fig. 8, the Sobol sensitivity analyses revealed differences primarily dependent on k . At high k , Sobol T and 1st indices for x_1 were dominant, indicating that the $\frac{BTEX}{T}RE$ was largely influenced by the x_1 (Fig. 9a and b); the Sobol T and 1st indices were 0.84 and 0.80 for Case I and 0.86 and 0.81 for Case II, respectively. The thickness of the LNAPL plume is thin and widely spread over the groundwater table at high k ; thus, $\frac{BTEX}{T}RE$ was highly influenced by the choice of screen depth in the

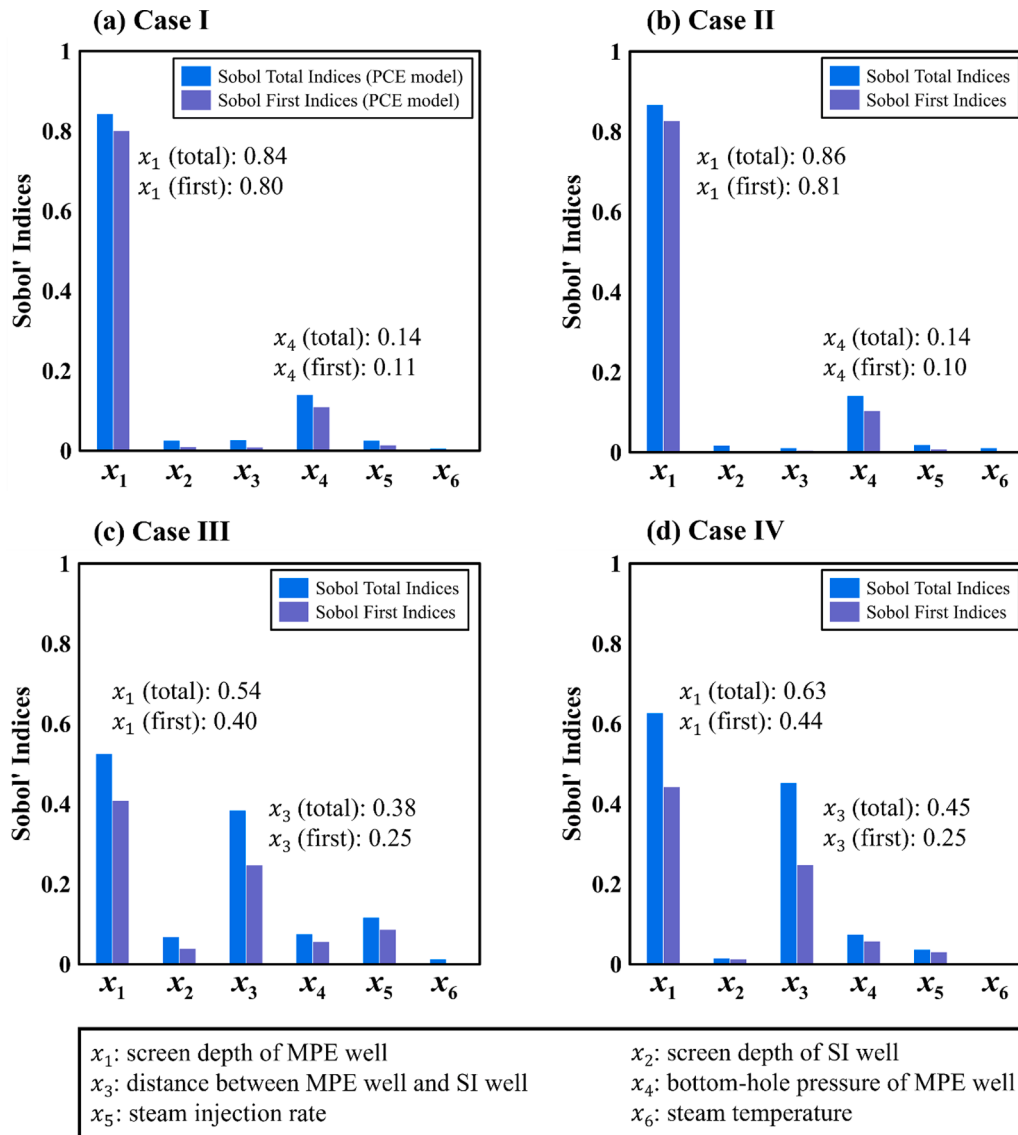


Fig. 9. Sobol total- and 1st-order indices for (a) Case I, (b) Case II, (c) Case III, and (d) Case IV.

MPE well. The 2nd influential factor was x_4 ; the Sobol T and 1st indices were 0.14 and 0.11 for Case I, and 0.14 and 0.10 for Case II, respectively. Even though the values for x_4 are significantly smaller than those for x_1 , their influences are still superior to the other input factors. This indicates that sufficient pressure gradient generated by the MPE well effectively enhances the $\frac{BTEX}{T}RE$ at high k . Finally, the Sobol indices of the other input factors (x_2 , x_3 , x_5 , x_6) related to the SI well were nearly zero.

At low k , the most sensitive factor was x_1 identical to high k , although the influence of x_1 was decreased compared to Cases I and II. The Sobol T and 1st indices of the x_1 were 0.54 and 0.40 for Case III, and 0.63 and 0.44 for Case IV, respectively (Fig. 9c and d). The difference between the Sobol T and 1st indices increased, indicating that the interaction between x_1 and the other factors (e.g. steam injection rate) became more important at low k . The 2nd influential factor was x_3 . The Sobol T and 1st indices of the x_3 were 0.38 and 0.25 for Case III, and 0.45 and 0.25 for Case IV, respectively. At low k , the spreading of the LNAPL plume was constrained, and consequently, the horizontal location and vertical depth of the MPE well became influential. Additionally, the influence of other input factors (x_2 and x_5) related to the SI wells increased. These results suggest that steam injection would be more effective for LNAPL remediation at the low k site. Finally, the change in x_6 within the designated range (i.e., 100–150 °C) did not affect the

$\frac{BTEX}{T}RE$, indicating that steam temperature above 100 °C is not cost-effective for remediating the BTEX.

4.3.2. Predicting LNAPLs remediation efficiency

The influences of 6 input factors on $\frac{BTEX}{T}RE$ was comprehensively quantified using the validated PCE models. A randomly generated 6 input factors determined a single prediction of $\frac{BTEX}{T}RE$ through the PCE model, and 10^5 Monte Carlo implementations revealed density plots of 6 input factors relating to empirical probability density functions of $\frac{BTEX}{T}RE$ (Fig. 10). The 6 density plots showing the correlation between 6 input factors (along the x-axis) and $\frac{BTEX}{T}RE$ (on the y-axis) are depicted for the 4 cases. The solid lines represent the linear fitting curves. Additionally, histograms and box-whisker diagrams were plotted to show the statistical distribution of $\frac{BTEX}{T}RE$.

Similar to Sobol global sensitivity, the PCE-based Monte Carlo prediction also differed significantly depending on the magnitude of k (Cases I and II vs. Cases III and IV). At high k (Cases I and II), the slopes of the fitting curves for the screen depth of the MPE well (x_1) were the steepest, confirming that x_1 had a strong effect on $\frac{BTEX}{T}RE$ (Fig. 10a and b). Interestingly, the bimodal distribution for x_1 , which is distinguished by the depth of the groundwater table, was shown. When the MPE well

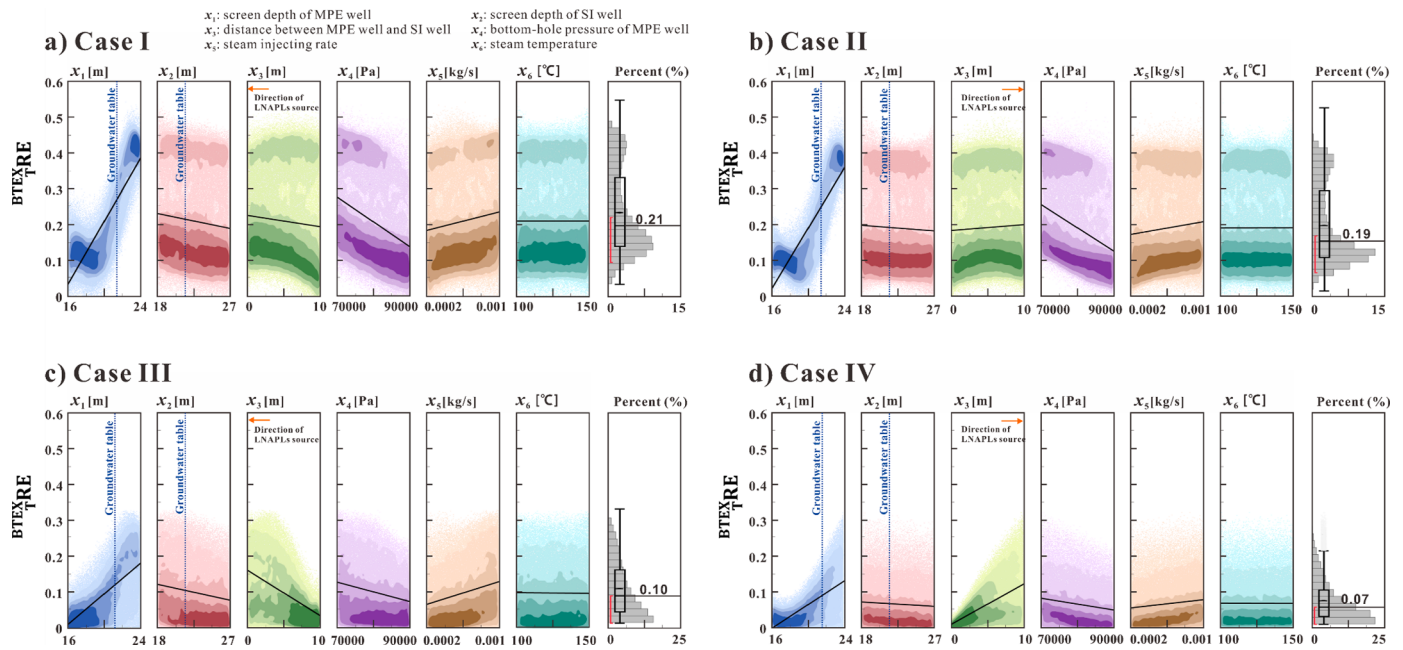


Fig. 10. The density plots were obtained by 10^5 implementations of the Monte-Carlo simulation based on the surrogate PCE models for (a) Case I, (b) Case II, (c) Case III, and (d) Case IV. In each density plot, predicted probabilistic distribution of the remediation efficiency according to the variation of the input factors were delineated. The statistical distributions were also provided as histograms and box-whisker plots at the rightmost of each figure.

was installed above the groundwater table (21.5 m), the average $\frac{BTRE}{T}$ was approximately 0.15. However, the $\frac{BTRE}{T}$ dramatically increased to over 0.4 when x_1 was slightly below the groundwater table; the $\frac{BTRE}{T}$ was maximized approximately 1.5 m below the groundwater table. This is attributed to the characteristic of LNAPLs floating above the groundwater table. When the MPE well was installed in the unsaturated zone, gaseous LNAPLs volatilized by steam injection were only removed. However, when the MPE well was located below the groundwater table, both liquid LNAPLs and LNAPLs dissolved in groundwater were removed effectively (Qi et al., 2020). In addition, the hydraulic head gradient developed by remediation wells accelerated the migration of the LNAPLs to the MPE well (Simon et al., 1999). Different from x_1 , the screen depth of the SI well (x_2) did not influence $\frac{BTRE}{T}$ significantly; the $\frac{BTRE}{T}$ slightly improved as x_2 was shallow. The distance between the MPE well and SI well (x_3) showed opposite slopes in Cases I and II because of the opposite direction of the LNAPL source (Fig. 7a and b); the $\frac{BTRE}{T}$ increased as x_3 was close to the LNAPL source. The 2nd largest input factor was the BHP of the MPE well (x_4). As x_4 approached the low limit (7×10^4 Pa), the average $\frac{BTRE}{T}$ increased. The effect of both the steam injection rate (x_5) and steam temperature (x_6) was insignificant for LNAPL removal. Finally, the histogram of $\frac{BTRE}{T}$ featured two peaks (high $\frac{BTRE}{T}$ and low $\frac{BTRE}{T}$), primarily split by the depth of the groundwater table.

At low k , the influence of x_1 was weaker than that of high k (Fig. 10c and d), but its influence was still the largest among the 6 input factors. Similar to high k , $\frac{BTRE}{T}$ increased when x_1 was located below the groundwater table. The x_3 was the 2nd important input factor; as x_3 was far from the LNAPLs source, the $\frac{BTRE}{T}$ approached zero. This indicates that the distance between the LNAPL source and remediation wells is critical at the low k field where LNAPLs migration is hindered. While the influence of x_3 became significant, the effect of x_4 decreased at low k . The effect of both x_5 and x_6 was small, similar to that for high k . Overall, at low k , the $\frac{BTRE}{T}$ were significantly smaller than high k , which is also reflected in the histograms and the box-whisker plots. This implies that the optimum choice of 6 input factors is more important for the high k .

5. Implications and limitations of PCE models

In this study, four surrogate PCE models accounting for LNAPLs remediation efficiency were developed for different conceptual cases (Case I, II, III, and IV). For each case, the number of training datasets to develop the PCE models was 150. However, the application of advanced adaptive-sparse PCE algorithm could substantially reduce the required number of the dataset by eliminating the non-influential interaction terms (Blatman and Sudret, 2011). Such computational advantages by reducing the number of training datasets would allow for developing surrogate PCE models representing more complex 3D heterogeneity numerical models with realistic geostructures and a large number of input factors such as hydrogeologic properties (e.g. permeability, porosity, water saturation), soil properties (e.g. particle-size distribution and capillary pressure), or geochemical properties (e.g. biodegradation coefficient, absorption coefficient). Additional to computational benefits, the surrogate PCE model can analytically link to the global sensitivity analysis and then be easily interconnected with other quantitative analyses (Fajraoui et al., 2011). For example, the global sensitivity analysis in this study revealed a crucial feature in LNAPLs remediation that the installation depth of the MPE well dominantly affected LNAPLs remediation efficiency. Subsequent Monte Carlo prediction further revealed the optimum depth of the MPE well that also supported the results obtained by numerical simulations in other previous studies (Qi et al., 2020).

Despite the aforementioned advantages of surrogate PCE models in LNAPLs remediation, researchers should be aware that the surrogate PCE model can not be a general or ultimate solution to interpret the LNAPLs transport behavior in the subsurface. Firstly, the surrogate PCE model developed is valid only in the conceptual model domain represented by numerical simulation. In the study, the numerical model was 2D and homogeneous. Thus, the surrogate PCE model should be reevaluated if researchers wish to investigate heterogeneous subsurface systems, 3D LNAPLs transport, or different target contaminants. Secondly, the surrogate PCE model solely depends on the internal computation results of the full-physics numerical LNAPLs transport model. Due to this reason, the choice of different numerical simulators may produce different PCE results. Accordingly, utilizing a more accurate and

effective surrogate PCE model should be accompanied by improvement of the underlying numerical model; For example, recent studies considered more realistic physical properties in simulating LNAPLs migration such as hysteresis effect of multi-phase fluids and also verified the LNAPLs model from the implementation of LNAPLs tank laboratory experiments (Pasha et al., 2014; Sookhak Lari et al., 2016).

6. Conclusion

Full-physics numerical simulations that honor multi-phase multi-component LNAPL transport and its remediation are challenging, primarily because of their extensive memory and CPU requirements. In addition to the complexities involved in geologic characterization and LNAPL transport, numerical simulations must be repeatedly performed to characterize uncertainties involved in input factors or to search for the optimum LNAPL removal efficiency. In this study, to overcome the challenges addressed in LNAPL transport and remediation, the surrogate PCE modeling technique was implemented in conjunction with the quantitative Sobol sensitivity analysis and Monte Carlo prediction.

The proposed workflow involves forward full-physics numerical LNAPLs transport modeling that generates a training dataset, development of a PCE-based surrogate model, global sensitivity analysis, and Monte Carlo prediction. For the full-physics numerical modeling, the TOUGH2-TMVOC was employed to simulate multi-phase and multi-component LNAPL transport. The conceptual model delineated the scenarios of the LNAPL-spilling, natural distribution, and remediation stages. Here, LNAPLs were assumed to be a mixture of BTEX components, allowing mass transfer (e.g., vaporization, condensation, and dissolution) among individual components. Finally, the removal efficiency of LNAPLs was assessed by implementing MPE and SI wells. The 4 cases, considering different permeabilities and locations of the SI well, were designed to develop PCE-based surrogate models. Within 4 cases, 6 factors related to well configuration were varied to assess the removal efficiency of LNAPLs.

By combining forward numerical modeling, global sensitivity analyses, and Monte Carlo prediction, the governing factors improving the efficiency of LNAPL remediation were identified. First, the screen depth of the MPE well was the most important among 6 factors associated with well configuration regardless of field permeability or location of the SI well. Because the LNAPL plume floated within the narrow capillary fringe above the groundwater table, the maximum remediation efficiency was predicted when the depth of the MPE well was slightly (1.5 m) below the groundwater table. Second, the permeability of contaminated sites evidently influenced the remediation conditions. At high k , both the screen depth and BHP of the MPE well are important, but the influence of the BHP of the MPE well was diminished at low k . Instead, the distance between the MPE and SI wells became important. Finally, at high k , the MPE well itself was enough to remove the LNAPLs, and thus, the configuration of the SI well was not as important. However, at low k where the mobility of LNAPLs was small, the SI well improved the remediation efficiency by volatilizing liquid LNAPLs to gaseous LNAPLs.

This study successfully demonstrated the capability of the PCE-based surrogate modeling for the quantitative analysis of LNAPL remediation, while overcoming the computational burden. In this study, the model was limited in 2D and homogeneous matrix to focus on the demonstration of the proposed research framework. As a future direction, however, it is expected that one can easily extend the proposed methodology to 3D heterogeneous aquifers with a consideration of various input factors. When extending the framework, the following elements should be accounted for. Results of the surrogate PCE model are site-specific and are subject to the performance of the chosen numerical simulator. In addition, it is important to specify proper minimum and maximum limits, and probability distribution of input factors based on target scenarios, such as target contaminants or hydrogeologic characteristics of the area.

Open research

The numerical simulation data used for evaluation of LNAPLs remediation efficiencies for the base case (Case I) in the study are available in Zenodo via <https://doi.org/10.5281/zenodo.5374335> (Kim and Han, 2021) with restricted access conditions.

CRediT authorship contribution statement

Taecheon Kim: Conceptualization, Data curation, Resources, Software, Formal analysis, Validation, Visualization, Methodology, Writing – original draft. **Weon Shik Han:** Conceptualization, Supervision, Resources, Writing – review & editing, Funding acquisition. **Jize Piao:** Methodology, Writing – review & editing. **Peter K. Kang:** Writing – review & editing. **Jehyun Shin:** Resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was supported by the Subsurface Environmental Management (SEM) project through the Korea Environmental Industry and Technology Institute (KEITI) funded by the Ministry of Environment (Grant Number: 2018002440003) and by Energy & Mineral Resources Development Association of Korea (EMRD) grant funded by the Korea government (MOTIE) (Data science based oil/gas exploration consortium). Peter K. Kang acknowledges the support by the National Science Foundation under Grant No. EAR-2046015. In addition, the authors would like to thank the useful discussion with Dowan Koo from Yonsei University, Republic of Korea.

References

- Adenekan, A., Patzek, T.W., Pruess, K., 1993. Modeling of multiphase transport of multicomponent organic contaminants and heat in the subsurface: numerical model formulation. *Water Resour. Res.* 29 (11), 3727–3740. <https://doi.org/10.1029/93WR01957>.
- Alizadeh, A., Piri, M., 2014. Three-phase flow in porous media: a review of experimental studies on relative permeability. *Rev. Geophys.* 52 (3), 468–521. <https://doi.org/10.1002/2013RG000433>.
- Anderson, W.G., 1987. Wettability literature survey-part 4: effects of wettability on capillary pressure. *J. Petrol. Technol.* 39 (10) <https://doi.org/10.2118/15271-PA>, 1,283–281,300.
- Asher, M.J., Croke, B.F., Jakeman, A.J., Peeters, L.J., 2015. A review of surrogate models and their application to groundwater modeling. *Water Resour. Res.* 51 (8), 5957–5973. <https://doi.org/10.1002/2015WR016967>.
- Baker, L., 1988. Three-phase relative permeability correlations. In: Paper presented at SPE Enhanced Oil Recovery Symposium. Tulsa, Oklahoma. Society of Petroleum Engineers.
- Bear, J., Ryzhik, V., Braester, C., Entov, V., 1996. On the movement of an LNAPL lens on the water table. *Transp. Porous Media* 25 (3), 283–311. <https://doi.org/10.1007/BF00140985>.
- Blatman, G., 2009. Adaptive Sparse Polynomial Chaos Expansions for Uncertainty Propagation and Sensitivity Analysis, (Doctoral Dissertation). Université Blaise Pascal, Clermont-Ferrand.
- Blatman, G., Sudret, B., 2008. Sparse polynomial chaos expansions and adaptive stochastic finite elements using a regression approach. *Comptes Rendus Mécanique* 336 (6), 518–523. <https://doi.org/10.1016/j.crme.2008.02.013>.
- Blatman, G., Sudret, B., 2010a. Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. *Reliab. Eng. Syst. Saf.* 95 (11), 1216–1229. <https://doi.org/10.1016/j.res.2010.06.015>.
- Blatman, G., Sudret, B., 2010b. An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Probab. Eng. Mech.* 25 (2), 183–197. <https://doi.org/10.1016/j.probengmech.2009.10.003>.
- Blatman, G., Sudret, B., 2011. Adaptive sparse polynomial chaos expansion based on least angle regression. *J. Comput. Phys.* 230 (6), 2345–2367. <https://doi.org/10.1016/j.jcp.2010.12.021>.
- Blunt, M.J., 2000. An empirical model for three-phase relative permeability. *SPE J.* 5 (04), 435–445. <https://doi.org/10.2118/67950-PA>.

- Bradford, S.A., Abriola, L.M., Leij, F.J., 1997. Wettability effects on two-and three-fluid relative permeabilities. *J. Contam. Hydrol.* 28 (1–2), 171–191. [https://doi.org/10.1016/S0169-7722\(97\)00024-7](https://doi.org/10.1016/S0169-7722(97)00024-7).
- Brennen, C.E., Brennen, C.E., 2005. *Fundamentals of Multiphase Flow*. Cambridge University Press, Pasadena, CA.
- Ciriello, V., de Barros, F.P., 2020. Characterizing the influence of multiple uncertainties on predictions of contaminant discharge in groundwater within a Lagrangian stochastic formulation. *Water Resour. Res.* 56 (10) <https://doi.org/10.1029/2020WR027867> e2020WR027867.
- Erning, K., Shafer, D., Dahmke, A., Luciano, A., Viotti, P., Petrangeli Papini, M., 2009. In: *Simulation of DNAPL infiltration into groundwater with differing flow velocities using TMVOC combined with PetraSim*, paper presented at Proc. TOUGH symposium 2009. Berkeley, California.
- Essaid, H.I., Bekins, B.A., Cozzarelli, I.M., 2015. Organic contaminant transport and fate in the subsurface: evolution of knowledge and understanding. *Water Resour. Res.* 51 (7), 4861–4902. <https://doi.org/10.1002/2015WR017121>.
- Fajraoui, N., Ramasomanana, F., Younes, A., Mara, T.A., Ackerer, P., Guadagnini, A., 2011. Use of global sensitivity analysis and polynomial chaos expansion for interpretation of nonreactive transport experiments in laboratory-scale porous media. *Water Resour. Res.* 47 (2) <https://doi.org/10.1029/2010WR009639>.
- Gupta, P.K., Yadav, B., Yadav, B.K., 2019. Assessment of LNAPL in subsurface under fluctuating groundwater table using 2D sand tank experiments. *J. Environ. Eng.* 145 (9), 04019048. [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0001560](https://doi.org/10.1061/(ASCE)EE.1943-7870.0001560).
- Guyant, E., Han, W.S., Kim, K.-Y., Park, E., Yun, S.-T., 2016. Leakage and pressurization risk assessment of CO₂ reservoirs: a metamodeling modeling approach. *Int. J. Greenhouse Gas Control* 54, 345–361. <https://doi.org/10.1016/j.ijggc.2016.10.004>.
- Hassanizadeh, S.M., Gray, W.G., 1993. Thermodynamic basis of capillary pressure in porous media. *Water Resour. Res.* 29 (10), 3389–3405. <https://doi.org/10.1029/93WR01495>.
- Huntley, D., Beckett, G., 2002. Persistence of LNAPL sources: relationship between risk reduction and LNAPL recovery. *J. Contam. Hydrol.* 59 (1–2), 3–26. [https://doi.org/10.1016/S0169-7722\(02\)00073-6](https://doi.org/10.1016/S0169-7722(02)00073-6).
- Jia, W., McPherson, B.J., Pan, F., Xiao, T., Bromhal, G., 2016. Probabilistic analysis of CO₂ storage mechanisms in a CO₂-EOR field using polynomial chaos expansion. *Int. J. Greenhouse Gas Control* 51, 218–229. <https://doi.org/10.1016/j.ijggc.2016.05.024>.
- Juanes, R., 2003. *Displacement Theory and Multiscale Numerical Modeling of Three-Phase Flow in Porous Media*. University of California, Berkeley.
- Kaluvarachi, J., Parker, J., 1990. Modeling multicomponent organic chemical transport in three-fluid-phase porous media. *J. Contam. Hydrol.* 5 (4), 349–374. [https://doi.org/10.1016/0169-7722\(90\)90025-C](https://doi.org/10.1016/0169-7722(90)90025-C).
- Kim, J., Corapcioglu, M.Y., 2003. Modeling dissolution and volatilization of LNAPL sources migrating on the groundwater table. *J. Contam. Hydrol.* 65 (1–2), 137–158. [https://doi.org/10.1016/S0169-7722\(02\)00105-5](https://doi.org/10.1016/S0169-7722(02)00105-5).
- Kim, T., Han, W.S., 2021. Numerical modeling for evaluation of LNAPLs remediation efficiency by configuration of wells. edited Zenodo. <https://doi.org/10.5281/zenodo.5374335>.
- Lenhard, R., Parker, J., 1987. Measurement and prediction of saturation-pressure relationships in three-phase porous media systems. *J. Contam. Hydrol.* 1 (4), 407–424. [https://doi.org/10.1016/0169-7722\(87\)90017-9](https://doi.org/10.1016/0169-7722(87)90017-9).
- Mayer, A.S., 2005. *Soil and Groundwater Contamination: Nonaqueous Phase Liquids*. American Geophysical Union, Washington, DC.
- McCray, J.E., Falta, R.W., 1997. Numerical simulation of air sparging for remediation of NAPL contamination. *Groundwater* 35 (1), 99–110. <https://doi.org/10.1111/j.1745-6584.1997.tb00065.x>.
- Mendoza, C.A., Frind, E.O., 1990. Advective-dispersive transport of dense organic vapors in the unsaturated zone: 1. Model development. *Water Resour. Res.* 26 (3), 379–387. <https://doi.org/10.1029/WR026i003p00379>.
- Miller, C.T., Christakos, G., Imhoff, P.T., McBride, J.F., Pedit, J.A., Trangenstein, J.A., 1998. Multiphase flow and transport modeling in heterogeneous porous media: challenges and approaches. *Adv. Water Resour.* 21 (2), 77–120. [https://doi.org/10.1016/S0309-1708\(96\)00036-X](https://doi.org/10.1016/S0309-1708(96)00036-X).
- Nielsen, D., Th. Van Genuchten, M., Biggar, J., 1986. Water flow and solute transport processes in the unsaturated zone. *Water Resour. Res.* 22 (9S), 89S–108S. <https://doi.org/10.1029/WR022i09Sp0089S>.
- Oliveira, L.L., Demond, A.H., 2003. Estimation of primary drainage three-phase relative permeability for organic liquid transport in the vadose zone. *J. Contam. Hydrol.* 66 (3–4), 261–285.
- Oostrom, M., Hofstee, C., Wietsma, T.W., 2006. Behavior of a viscous LNAPL under variable water table conditions. *Soil Sediment Contam.* 15 (6), 543–564. <https://doi.org/10.1080/15320380600958976>.
- Pasha, A.Y., Hu, L., Meegoda, J.N., 2014. Numerical simulations of a light nonaqueous phase liquid (LNAPL) movement in variably saturated soils with capillary hysteresis. *Can. Geotech. J.* 51 (9), 1046–1062. <https://doi.org/10.1139/cgj-2012-0165>.
- Piao, J., Han, W.S., Kang, P.K., Min, B., Kim, K.-Y., Han, G., Park, J.G., 2020. A hybrid optimization methodology identifying optimal operating conditions for carbon dioxide injection in geologic carbon sequestration. *Int. J. Greenhouse Gas Control* 98, 103067. <https://doi.org/10.1016/j.ijggc.2020.103067>.
- Pruess, K., 2003. Numerical simulation of CO₂ leakage from a geologic disposal reservoir including transitions from super- to sub-critical conditions, and boiling of liquid of CO₂. *SPE J.* 9 (02), 237–248. <https://doi.org/10.2118/86098-PA>.
- Pruess, K., Battistelli, A., 2002. *TMVOC, a Numerical Simulator for Three-Phase Non-Isothermal Flows of Multicomponent Hydrocarbon Mixtures in Saturated-Unsaturated Heterogeneous Media* (Report LBNL-49375). Lawrence Berkeley National Laboratory. Retrieved from, Berkeley, CA. <https://escholarship.org/uc/it-em/73m7d7c4>. Google Scholar.
- Qi, S., Luo, J., O'Connor, D., Wang, Y., Hou, D., 2020. A numerical model to optimize LNAPL remediation by multi-phase extraction. *Sci. Total Environ.* 718, 137309. <https://doi.org/10.1016/j.scitotenv.2020.137309>.
- Rajabi, M.M., Ataie-Ashtiani, B., Simmons, C.T., 2015. Polynomial chaos expansions for uncertainty propagation and moment independent sensitivity analysis of seawater intrusion simulations. *J. Hydrol. (Amst.)* 520, 101–122. <https://doi.org/10.1016/j.jhydrol.2014.11.020>.
- Razavi, S., Tolson, B.A., Burn, D.H., 2012. Review of surrogate modeling in water resources. *Water Resour. Res.* 48 (7) <https://doi.org/10.1029/2011WR011527>.
- Reid, R.C., Prausnitz, J.M., Poling, B.E., 1987. *The Properties of Gases and Liquids*. McGraw Hill Book Co, New York, NY.
- Robin, M., Gillham, R., 1987. Field evaluation of well purging procedures. *Groundwater Monitor. Remediat.* 7 (4), 85–93. <https://doi.org/10.1111/j.1745-6592.1987.tb00967.x>.
- Rogers, S.W., Ong, S.K., 2000. Influence of porous media, airflow rate, and air channel spacing on benzene NAPL removal during air sparging. *Environ. Sci. Technol.* 34 (5), 764–770. <https://doi.org/10.1021/es990112>.
- Russell, T.F., 1995. Modeling of multiphase multicontaminant transport in the subsurface. *Rev. Geophys.* 33 (S2), 1035–1047. <https://doi.org/10.1029/95RG00289>.
- Scholes, O.N., Clayton, S., Hoadley, A.F.A., Tiu, C., 2007. Permeability anisotropy due to consolidation of compressible porous media. *Transp. Porous. Media* 68 (3), 365–387. <https://doi.org/10.1007/s11242-006-9048-5>.
- Schroth, M., Istok, J.D., Selker, J.S., Oostrom, M., White, M.D., 1998. Multifluid flow in bedded porous media: laboratory experiments and numerical simulations. *Adv. Water Resour.* 22 (2), 169–183. [https://doi.org/10.1016/S0309-1708\(97\)00043-2](https://doi.org/10.1016/S0309-1708(97)00043-2).
- Simon, M., Saddington, B., Zahraleslamzadeh, Z., DiLorenzo, J., 1999. *Multi-Phase Extraction: State-of-the-Practice*. Tetra Tech Environmental Management, Vienna, VA. Retrieved from. <https://apps.dtic.mil/sti/citations/ADA487281>.
- Song, J., Yang, Y., Chen, G., Sun, X., Lin, J., Wu, J., Wu, J., 2019. Surrogate assisted multi-objective robust optimization for groundwater monitoring network design. *J. Hydrol. (Amst.)* 577, 123994. <https://doi.org/10.1016/j.jhydrol.2019.123994>.
- Sookhak Lari, K., Davis, G.B., Johnston, C.D., 2016. Incorporating hysteresis in a multi-phase multi-component NAPL modelling framework; a multi-component LNAPL gasoline example. *Adv. Water Resour.* 96, 190–201. <https://doi.org/10.1016/j.advwatres.2016.07.012>.
- Sookhak Lari, K., Johnston, C.D., Rayner, J.L., Davis, G.B., 2018. Field-scale multi-phase LNAPL remediation: validating a new computational framework against sequential field pilot trials. *J. Hazard. Mater.* 345, 87–96. <https://doi.org/10.1016/j.jhazmat.2017.11.006>.
- Stone, H., 1970. Probability model for estimating three-phase relative permeability. *J. Petrol. Technol.* 22 (02), 214–218. <https://doi.org/10.2118/2116-PA>.
- Sudret, B., 2008. Global sensitivity analysis using polynomial chaos expansions. *Reliab. Eng. Syst. Saf.* 93 (7), 964–979. <https://doi.org/10.1016/j.res.2007.04.002>.
- Viana, F.A., 2013. In: *Things you wanted to know about the Latin hypercube design and were afraid to ask*, paper presented at 10th World Congress on Structural and Multidisciplinary Optimization. Orlando, Florida.
- Wipfler, E., Ness, M., Breedveld, G., Marsman, A., Van Der Zee, S., 2004. Infiltration and redistribution of LNAPL into unsaturated layered porous media. *J. Contam. Hydrol.* 71 (1–4), 47–66. <https://doi.org/10.1016/j.jconhyd.2003.09.004>.
- Wu, Y.-S., Forsyth, P.A., 2001. On the selection of primary variables in numerical formulation for modeling multiphase flow in porous media. *J. Contam. Hydrol.* 48 (3–4), 277–304. [https://doi.org/10.1016/S0169-7722\(00\)00180-7](https://doi.org/10.1016/S0169-7722(00)00180-7).
- Xing, Z., Qu, R., Zhao, Y., Fu, Q., Ji, Y., Lu, W., 2019. Identifying the release history of a groundwater contaminant source based on an ensemble surrogate model. *J. Hydrol. (Amst.)* 572, 501–516. <https://doi.org/10.1016/j.jhydrol.2019.03.020>.
- Xiu, D., 2010. *Numerical Methods for Stochastic Computations*. Princeton university press, Princeton, NJ.
- Yang, Y., Li, J., Xi, B., Wang, Y., Tang, J., Wang, Y., Zhao, C., 2017. Modeling BTEX migration with soil vapor extraction remediation under low-temperature conditions. *J. Environ. Manage.* 203, 114–122. <https://doi.org/10.1016/j.jenvman.2017.07.068>.
- Zhao, Y., Qu, R., Xing, Z., Lu, W., 2020. Identifying groundwater contaminant sources based on a KELM surrogate model together with four heuristic optimization algorithms. *Adv. Water Resour.* 138, 103540. <https://doi.org/10.1016/j.advwatres.2020.103540>.