An Artificial Intelligence Based Self-Adaptive Dynamic Process Control System for Enhancing In-Situ Bioremediation of Benzene-Contaminated Groundwater

Enter authors here: L. He[1*†], X. Duan[1†], C. Li[1], Y. Xu[1], and M. He[1]

[1]State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, Tianjin 300350, China

[*]Corresponding author: Li He (helix111@tju.edu.cn)

[†]L. He and X. Duan contributed equally to this manuscript

Key Points:

- Perform a simulation of enhanced remediation process shown by a contaminant fate and transport model
- Develop a self-adaptive dynamic process control (SADPC) system and provide suggestion for guidelines and policies on SADPC systems
- Improve the removal rate of benzene in groundwater during the bioremediation process

Abstract

This study develops a new AI-based Self-Adaptive DPC (SADPC) system based on stepwise inference combing with genetic algorithm optimization technologies, including a filtered-clustering inference prediction model (FCI simulator), a stepwise inference controller (SI emulator), a model predictive control controller (MPC controller), a 1st-stage optimizer, and a 2nd-stage optimizer. This system effectively reflects the dynamics and complexity of the biodegradation process and realizes the control for the remediation system based on the feedback information. To achieve this goal, a statistical model for simulating the bioremediation process through the FCI simulator is proposed, which can predict the resulting contamination situation based on the previous contamination situation and control action. Then a bridge between control actions and contamination situations is established through the SI emulator, which can generate a control action based on a given contamination situation. Through running the SADPC system, the desired control action can be identified. Results show that The SADPC system increases the removal rate of benzene and arrives at the remediation goal earlier than other systems. This suggested decision makers that guidelines and policies on remediation-oriented SADPC systems could be tentatively investigated, developed, and applied in the future effort.

Keywords

Self-adaptive dynamic process control, in-situ bioremediation, contaminant fate and transport modeling, physically groundwater simulation

1 Introduction

In-situ bioremediation (ISB) techniques (Albers et al., 2015; Zhang et al., 2020) aim to enhance the biodegradation of organic constituents in the subsurface by encouraging the growth and reproduction of indigenous microorganisms. The ISB technique involves a mechanism for stimulating and maintaining the activity of intrinsic bioremediation processes, by which indigenous microbes convert contaminants to innocuous end products via electron acceptor and/or inorganic nutrient amendments. The normal operation of ISB consists of routine checking of operation and maintenance of equipment, groundwater levels, extraction and injection rates, groundwater electron acceptor concentrations, nutrient levels, pH, and conductivity.

System optimization approaches have been demonstrated to be useful in enhancing remediation efficiency and reduce remediation cost during water treatment and remediation (Chiandussi et al., 2012; He et al., 2008a; He et al., 2008b; He et al., 2008c; Passino, 2002; Sun et al., 2020; Wang et al., 2020). Compared with system optimization approaches, dynamic process control (DPC) could be a better way in fulfilling real-time system optimization by temporally regulating a set of operating conditions such as additions of electron acceptors and nutrients, groundwater extraction and injection rates, remedial cleanup time, etc. Various studies have been undertaken on development and applications of DPC techniques (Ahmed & Rodriguez, 2020; Bashivan et al., 2019; Bechet et al., 2016; Diangelakis et al., 2016; Liu et al., 2016; Mayne, 2014; Miller et al., 2016; Stentoft et al., 2021; Zeng & Liu, 2015). For example, Stentoft et al. (2021) proposed a general model predictive control algorithm to achieve the optimal operating conditions by controlling the effluent concentrations, total costs, and other management objectives. This approach allows the water resource recovery facilities to quickly accommodate new control requirements. Liu et al. (2016) presented an Event-driven Model Predictive Control (EMPC) method to ensure that the flows of sewage streams containing the dosed chemical are reasonably distributed throughout the sewer networks. The EMPC strategy substantially enhanced the performance of sulfide mitigation when dealing with the corrosion and odor problems. Ahmed & Rodriguez (2020) demonstrated a non-linear model predictive control (NMPC) system to optimize the automatic start-up of anaerobic digesters, which achieved a higher target methane production rate and superior control variables set-point tracking error performance.

However, three challenges of the conventional DPC techniques lead to the difficulty in applying to a general ISB system. First, conventional DPC depends on the use of a set of nonlinear state equations group (or prediction model) representing the input-output relations. While there is difficulty in analytically or numerically solving the equations group, the DPC would fail to work because of extremely low solution efficiency. Because the ISB prediction model is computationally costly, proxy modeling may be a good means of solving this challenge, i.e., to produce a set of proxy models to replace initial ones through statistical or artificial intelligence (AI) methods. Usually, proxy models have the advantages of computation-rapid, result-stable, and error-tolerable (Gopalakrishnan et al., 2011; Gorelick and Zheng 2015; He et al., 2008a; He et al., 2008b;Meray et al., 2022; Siade et al., 2020; Stramer et al., 2010). Second, generation of optimal operating conditions within a given time period by conventional DPC relies on the difference (or error) between the predicted remediation performance and pre-determined level. For facilitating computation, a prediction model implied in the DPC framework is generally assumed to maintain static (without any variation) during the entire remediation process. This assumption may not be suitable particularly where complex hydrogeological conditions and biochemical process exist in the groundwater. A feasible approach is to introduce self-adaptive prediction or proxy models that can be dynamically trained and improved subject to external environmental variations. Third, it tends to make predictions that cannot meet expectations when dealing with the complicated situation of contaminant degradation, since the conventional DPC prediction model remains static during the entire remediation process. To alleviate the problem of falling into the local optimality dilemma triggered by the DPC static prediction model, a near-ideal biodegradation process can be obtained by conducting a second-stage optimization and developing the predicted trajectory (setpoint curve) based on the entire process prediction of the DPC system.

Therefore, this paper aims to present a new AI-based self-adaptive DPC system (SADPC) for enhancing in-site bioremediation of benzene-contaminated groundwater due to non-aqueous phase liquids (NAPLs) leakage from underground storage tank. The DPC system includes a filtered-clustering inference prediction model (also called FCI simulator) (Gerber & Horenko, 2015; Pizzagalli et al., 2019; Zhan et al., 2018), a stepwise inference controller (SI emulator), an 1st-stage optimizer, a model predictive control controller (MPC controller), and a biodegradation process. The MPC controller includes an FCI simulator, an 2nd-stage optimizer, and an error regulator.

This task entails: 1) developing a statistical model for simulating the bioremediation process through the FCI simulator, which can predict the resulting contamination situation based on the previous contamination situation and control action; 2) establishing a bridge between control actions and contamination situations through the SI emulator, which can generate a control action based on a given contamination situation; 3) running the SADPC system to identify the desired control action.

2 Materials and Methods

2.1 Development of the pilot-scale reactor

A pilot-scale reactor was developed and used to physically simulate the flow and transport of benzene (gasoline) in the groundwater (He, 2008; He et al., 2008a). It also facilitated the implementation of enhanced in-situ biodegradation and the relevant simulation efforts. The reactor (Figure S1) is of cuboid shape with an interior dimension of Length × Width × Height = 3.6 × 1.2 × 1.4 m3. It was composed of four sections, each of which contained a supporting part, a loading manhole, and two observation windows. More details regarding the reactor were shown in the supporting information.

For the simulation of hydrocarbon leakage, 12 liters of gasoline were injected into the bottom of the second soil layer at an upper stream location during a 1.5-day period. At the same time, tap water from a water container was pumped into the system as groundwater inflow at a rate of 20 L/day (through a peristaltic pump). The water level in the upstream gauge was 55 cm high and that in the downstream one was 45 cm high. After the leakage period, such flow conditions were maintained for 40 days to simulate the process of natural attenuation in the subsurface. The enhanced in-situ biodegradation process was then started right after this 40-day period. The experiment of flow and transport lasted 40 days after the gasoline leakage, followed by a 22-day enhanced in-situ biodegradation action. Environmental managers are more concerned with benzene than toluene, ethylbenzene, and xylenes (TEX) due to the fact that benzene is highly toxic and carcinogenic. In addition, during the remediation process, concentrations of TEX would become much lower than the respective environmental criteria as long as the benzene concentration is lower than the regulated criterion. Therefore, only benzene concentrations were analyzed in this study. The set-up of the reactor and the detailed analysis of the pilot-scale experimentation can be seen in Sections S1 and S2 of the supporting information (McDonald & Harbaugh, 1988; Jimenez et al., 2006; Zhang et al., 2008; Wolicka et al., 2009; Liang et al., 2013; Niswonger & Prudic, 2013; Xin et al., 2013; Yang et al., 2019; Hu et al., 2021; Umar et al., 2021).

2.2 Contaminant fate and transport modeling in the groundwater

A critical step in understanding the impact of a subsurface release of NAPL is a modeling analysis of the NAPL flow and transport and fate of its crucial constituents. The 3D multiphase and multicomponent (3DMM) model is used to simulate contaminant fate and transport in the groundwater. The basic mass conservation equation for components in the subsurface can be written as follows (Li et al., 2007; Schaerlaekens et al., 2005):

$$\frac{\partial}{\partial t}(\phi \tilde{C}_k \rho_k) + \vec{\nabla} \cdot [\sum_{l=1}^{n_p} \rho_k (C_{kl} \vec{u}_l - \phi S_l \vec{\vec{D}}_{kl} \cdot \vec{\nabla} C_{kl})] = R_k \quad (1)$$

where k is the component index; l is the phase index; ϕ is the soil porosity; $\tilde{C}_k$ is the overall concentration of component k (volume fraction); ρ_k is the density of component k [ML^{-3}]; n_p is the number of phases; C_{kl} is the concentration of component k in phase l (volume fraction); $\vec{u}_l$ is the Darcy velocity of phase l [LT^{-1}]; S_l is the saturation of phase l; R_k is the total source/sink term for component k (volume of the component k per unit volume of porous media per unit time); $\vec{\vec{D}}_{kl}$ is the dispersion tensor. The overall concentration ($\tilde{C}_k$) denotes the volume of the component k summed over all phases. Formulas, solution methods and other details are given in the Section S3 in the supporting information (Bear, 1979; Faust et al., 1989; Delshad et al., 1996).

The model can be solved numerically through the block-centered finite difference method, and it is possible to obtain the concentration of specified component k in phase l (C_{kl}) at a certain time. In addition, the biodegradation model with single substrate, single electron acceptor and single biological species should be required for a system (de Blanc, 1998; Huang et al., 2006). The solution to the flow equations was used as the initial conditions for the biodegradation reactions. By incorporating the component concentrations obtained through the pollutant migration model (C_{kl}) into the biodegradation modeling of contaminants in the groundwater, the substrate degradation rate during this time period can be computed. Details regarding the biodegradation modeling of contaminants in the groundwater are shown in Sections S4 and S5 (Rittmann et al., 1991; Chang & Alvarez-Cohen, 1995; de Blanc, 1998; Chang & Alvarez-Cohen, 2010).

2.3 Framework of the Study Method

In this paper, an artificial intelligence-based self-adaptive dynamic process control (SADPC) system for enhancing in-situ bioremediation of benzene-contaminated groundwater is established. SADPC is used to temporarily adjust a set of operating conditions in the aforementioned biodegradation process to achieve real-time system optimization. For the realization of in-situ bioremediation of benzene-contaminated groundwater, four groundwater control models (i.e., SI emulator, FCI optimizer, DPC system and SADPC system) were simultaneously used to predict and adjust operating conditions for efficient pollutant degradation.

Considering the high complexities and dynamics of the bioremediation system, it is inevitable that some important information might be missed/ignored when establishing a biodegradation model since almost all the models are a selective, dynamic abstraction of reality. In some situations, if there are a large number of experimental data, a statistical relationship can be developed to substitute the general simplified model (Huang et al., 2006).

In this study, a set of surrogate simulators can be established to quantify the relationship between pumping/injecting flow rate and benzene concentration by employing a stepwise cluster analysis (SCA) method, detailed descriptions have been shown in He et al (He, 2008; He et al., 2008a; He et al., 2008b). More information can be seen in the supporting information (Section S6) (Rao, 1952). To determine the optimal repair strategy, an FCI simulator was presented based on the SCA method (Zou et al., 2009).

In the FCI simulator, the relationships between contaminant concentrations and remediation operating conditions can be established through the filtered-clustering inference method based on a number of simulation runs. Given the pollution situation, the optimal operating conditions of the FCI simulator can be obtained under the constraints of the optimization objective. Based on the FCI simulator, the FCI optimizer was developed to optimize the biodegradation process (the framework is shown in Figure S2). Sections S7 and S8 of the supporting information details the procedures of the FCI simulator and the optimization model for the FCI optimizer (Maybeck, 1979; Jacobs, 1993).

After the optimal operation conditions for each scenario are determined, the SI emulator is developed through the obtained knowledge base. For the SI emulator, the corresponding operating conditions can be obtained by a given benzene concentration, and the benzene concentration of the next stage can be obtained through the biodegradation process. The framework of the SI emulator is shown in Figure S3. The operating conditions coming from the SI emulator cannot be considered the optimal one because there are no standard optimization curves for the biodegradation process as a reference for the experiments. Therefore, operating conditions must be optimized before it is applied to the real bioremediation process, which can be realized through the adjustment of the ranges of control conditions.

At the same time, since the operating conditions obtained by the FCI simulator are optimized for cost minimization without taking remediation efficiency into consideration, the actual bioremediation often does not achieve ideal results and cannot be used as an actual optimization curve to guide in situ bioremediation of groundwater system. Therefore, the minimized operation cost and maximized degradation efficiency should be considered. For the operating conditions, decisions of oxygen and nutrient injection rates and groundwater extraction rates directly affect the operating cost. The lower the injection or extraction rate, the lower the cost and contaminant removal rate. According to the content above, the optimization model for the DPC system in this study is given in the Section S9 in the supporting information (Huang et al., 2008).

For the poor performance of traditional DPC technology applied to general ISB systems, this study proposes a SADPC technology to reduce the impact of the defects of traditional DPC technology itself. According to this improvement, the reference trajectory (i.e., prediction curve/setpoint curve) can be obtained based on the prediction results from the entire DPC process. Then rolling optimization is performed to update the prediction curve in the next time period. The optimization model for the SADPC system in this study is given in Section S10. GA is used to solve all the developed discrete and nonlinear model. More information on GA can be seen in the supporting information (Section S11) (Holland, 1975; Kuo et al., 2006; Matott et al., 2006; Stramer et al., 2010; Opher & Ostfeld, 2011; Greenland et al., 2016; Hou et al., 2017; Liu et al., 2017; Shen et al., 2018; Liao et al., 2020).

The framework of the DPC system is presented in Figure 1 (a). Based on the DPC system, the SADPC system can be optimized by adding an MPC controller (Figure 1 (b)). The major components include a SI emulator, an FCI simulator, an MPC controller, and an optimization procedure1. In Figure 1, $X(t)$ is the input for the SI emulator, the FCI simulator and the MPC controller; $X_P(t+1)$ is the output of the FCI simulator; $X_r(t+1)$ is the setpoint; $e(t+1)$ is the error between $X_P(t+1)$ and $X_r(t+1)$; $X^*(t+1)$ is the optimal contamination situation after system operation; $U(t)$ is the control action coming from the SI emulator; $U'(t)$ is the tentative control signal; $U^{1*}(t)$ is the optimal control action after the 1st-stage optimization procedure; $U^{2*}(t)$ is the optimal control action after the 2nd-stage optimization procedure. The specific operation program of the MPC controller can be seen in Section S12 in the supporting information.

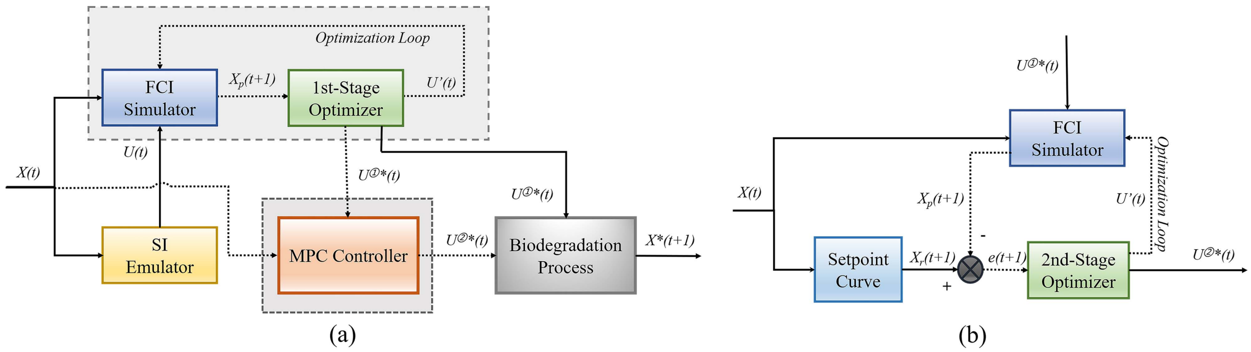


Figure 1. Framework of the SADPC system(a) and MPC controller (b)

Figure 2 shows the framework of the study method. The general procedure of developing a process control system for enhanced in-situ biodegradation consists of eight steps. The specific steps can be seen in Section S13 in the supporting information. Given the same initial benzene concentration, the predicted optimal degradation strategies of these four developed groundwater control models are various. By comparing the restoration processes and results, the optimal groundwater control models and operation strategies can be decided through the comparison of their degradation processes and the removal results.

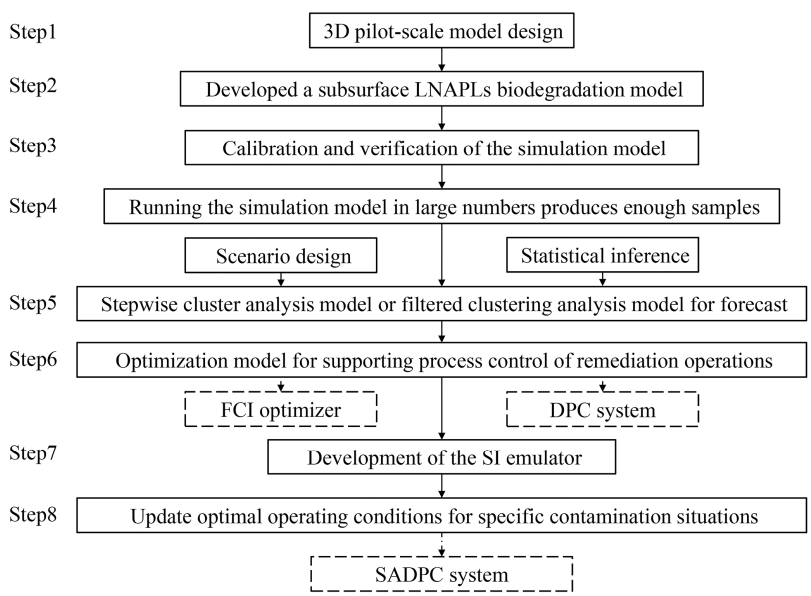


Figure 2. Flowchart of the solution method

3 Results

3.1 Clustering analysis

The developed NAPLs biodegradation model can then be used for simulating the system's responses under various operating conditions. However, it will bring high complexities and computation requirements if the developed simulation model is directly incorporated into the optimization framework. The filtered clustering method can be used to establish the relationship between the remediation efforts (i.e., pumping/injecting rates of selected wells) and the system's responses (i.e., benzene concentrations). The developed simulation model was used to generate a large number of inputs and outputs for supporting the establishment of such a relationship.

According to the characteristics of the soil profile, the NAPLs fate and transport and the contaminant plume movement, benzene concentrations in six wells were used as the representatives of the contamination situation in the groundwater. These included wells 5, 7, 8, 10, 11 and 12 (Benzene concentrations in these wells were denoted as $x_1^0, x_2^0, x_3^0, x_4^0, x_5^0$, and x_6^0). In order to reflect as many contamination situations as possible, a large range of the benzene concentration levels was considered. The maximum benzene concentration was 30 mg/L, and the minimum was 0 mg/L. Within this range, 50 concentration levels were generated randomly for each concerned well such that 50 contamination situations were produced (Table S5) (Lenczewski et al., 2003; Zhang et al., 2020).

A groundwater pumping system was used to circulate nutrients and oxygen through the contaminated aquifer. The process involves (a) the introduction of aerated and nutrient- and biomass-enriched water into the contaminated zone through two injection wells, and (b) the recovery of the down-gradient water through two extraction wells. The amendments were circulated through the contaminated zone to provide mixing and intimate contacts among the oxygen, nutrients, contaminant, and microorganisms. Therefore, the pumping/injecting rates directly affected the contaminant removal efficiency and system operation cost. In this study, pumping/injecting rates of selected wells were identified as the main control conditions.

The ranges of pumping/injecting rates were determined by considering the soil porosities and permeabilities in the pilot system and testifying them through the developed biodegradation model. The maximum flow rate was set as 40 L/day while the minimum was 10 L/day. The biomass/oxygen/nutrient concentrations were 20/8/1500 (mg/L) in the injecting fluid, respectively. Totally 50 scenarios of the operating conditions were randomly generated (Table S6 in the supporting information). The relevant control variables were denoted as u_1 (injection rate for well I, L/d), u_2 (injection rate for well II, L/d), u_3 (extraction rate in well III, L/d), and u_4 (extraction rate in well IV, L/d).

The combination of the 50 contamination-level scenarios and the 50 operating-condition scenarios led to 2500 scenarios. Correspondingly, 2500 input files were produced for the developed NAPLs biodegradation model. The experimental results indicated that benzene concentrations in the groundwater reduced significantly 18 days after the remediation started. Therefore, a 22-day duration was set, which was divided into 11 2-day periods. For each contamination-level scenario ($x_1^0, x_2^0, x_3^0, x_4^0, x_5^0$, and x_6^0), 50 sets of data about (1) the respondent percentage of benzene mass removal (η) and (2) the operating conditions of enhanced in-situ biodegradation (u_1, u_2, u_3, and u_4) can be obtained from simulation runs. Basing on the 300 cluster trees obtained through the filtered clustering analysis in total, the value of η can be predicted given the inputs of operating conditions.

The relationship between the process operating conditions and respondent value of η under 50 contamination-level scenarios were established through the filtered clustering analysis based on a large number of simulation runs under 50 operating-condition scenarios. For each contamination-level scenario (with the initial benzene concentrations of ($x_1^0, x_2^0, x_3^0, x_4^0, x_5^0$, and x_6^0), the resulting cluster tree system can be incorporated into a discrete and nonlinear optimization model. Genetic algorithm was used to solve the developed discrete and nonlinear model under each contamination-level scenario[38-41]. The number of generations was set as 200; the crossover rate (R^{CRO}) was 0.6; the mutation rate (R^{MUT}) was 0.003; and the number of the initial population was 70.

Over-parameterization is described as the scenario where the number of parameters of the model is redundant compared to the training dataset. Its high power consumption and memory occupation can degrade the performance of the model and make prediction action worse. This research is not over-parameterized because the network pruning method (Akyol, 2020; Hao & Chiang, 2006; Itoh & Adachi, 2017; Sun et al., 2015) was used to alleviate the problems and achieve the best results by evaluating the importance of the parameters based on the absolute values and removing the unimportant parameters. Then the regularizer can be added to the loss in order to make the weights sparse in the training process. During the experiment, a number of attempts was made to make sure that the model performs well under these specific model parameters.

3.2 Process control action analysis

Process control is used for operating the ISB system based on the SI emulator, FCI simulator and GA-based optimizer. In the DPC system, firstly, benzene concentrations at the concerned wells at the beginning of time period t were monitored. Then the highest contaminant concentration anywhere in the mesh (Ƀ$^{\mathrm{MAX}}$) and the percentage of benzene mass removal (η) can be used as inputs for the SI emulator to generate the optimal operating schemes correspondingly for the period t. the outputs are pumping/injecting rates of wells I, II, III, and IV. For the FCI simulator, the inputs include operating conditions of selected wells and the corresponding benzene concentrations; the outputs are the highest contaminant concentration anywhere in the mesh (Ƀ$^{\mathrm{MAX}}$) and the percentage of benzene mass removal (η) (detailed in Table S7 in the supporting information). It means that the pumping/injecting rates depends on the biodegradation process of benzene at time period t. Next, benzene concentrations in the concerned wells were monitored at the end of period t; then they can be regarded as new initial states for the next time period. The entire biodegradation process in benzene-contaminated groundwater could be controlled with cost-effective operational decisions step by step.

According to the degradation situation and operation process, a second-stage rolling optimization model (SADPC system) was used to meet the further expectation. An ideal setpoint curve can be produced based on the control process and benzene removal information of DPC system and new setpoint curves should be updated in the next optimization period. Therefore, a SADPC system can improve the biodegradation performance in the DPC system.

The contaminant concentration distribution of Day 57 was used as initial conditions for the SADPC system, and the pumping/injecting rates were assumed to be adjusted every two days. Figure 3 presents the 11 optimal operating conditions for the 11 2-day periods of the entire remediation duration in the DPC system. Results indicate that the operating conditions of 4 wells varied significantly under different instructions. Both highest injection rates for well I (u1) and well II (u2) are found on Day 6 at 0.022 and 0.025 m^3/d respectively, and the extraction rates for well III (u3) and IV (u4) reach the highest on Day 10 with 0.034 and 0.038 m^3/d respectively.

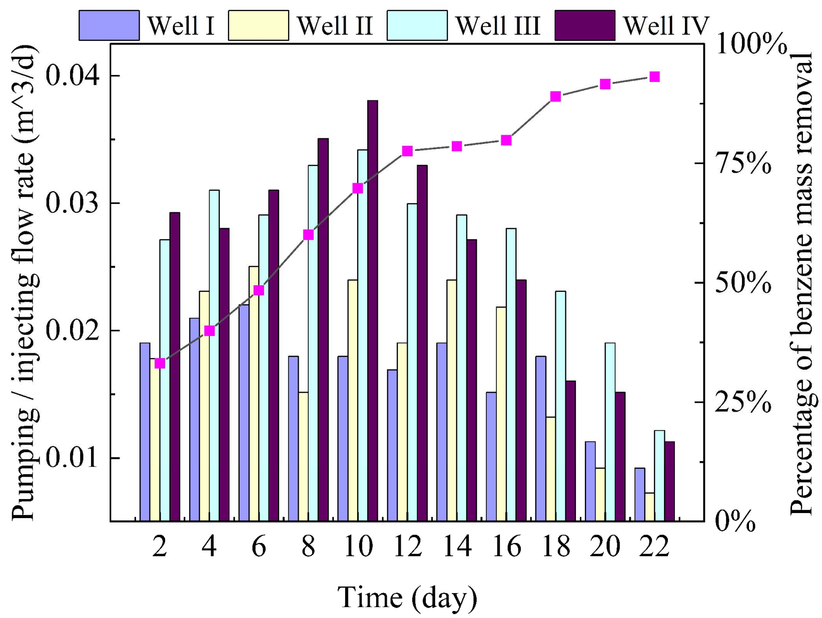


Figure 3. Optimal pumping/injecting rates of Well I to IV for the remediation duration and the corresponding percentage of benzene mass removal

Figure 4 presents the predicted remediation results of the DPC and SADPC systems from Day 2 to Day 22 after the leakage. It is shown that the contamination level has been reduced significantly through both systems. The benzene concentrations are over 0.5 mg/L at the initial stage and then transport and decrease gradually with groundwater flow. Therefore, the peak benzene concentrations at the upstream are gradually getting decreased over time to only 0.2 mg/L on day 22 in the DPC system, which has been removed around 60% of the initial contamination. The SADPC system focus on the biodegradation process from Day 12 to Day 22, so Figure 4 also shows the different levels between the DPC system and the SADPC system. The levels of benzene dispersion through the SADPC system are relatively lower compared with the DPC system since Day 14, and reach the degradation goal on Day 20. It indicates that the SADPC system has outstanding performance at benzene removal in the groundwater system.

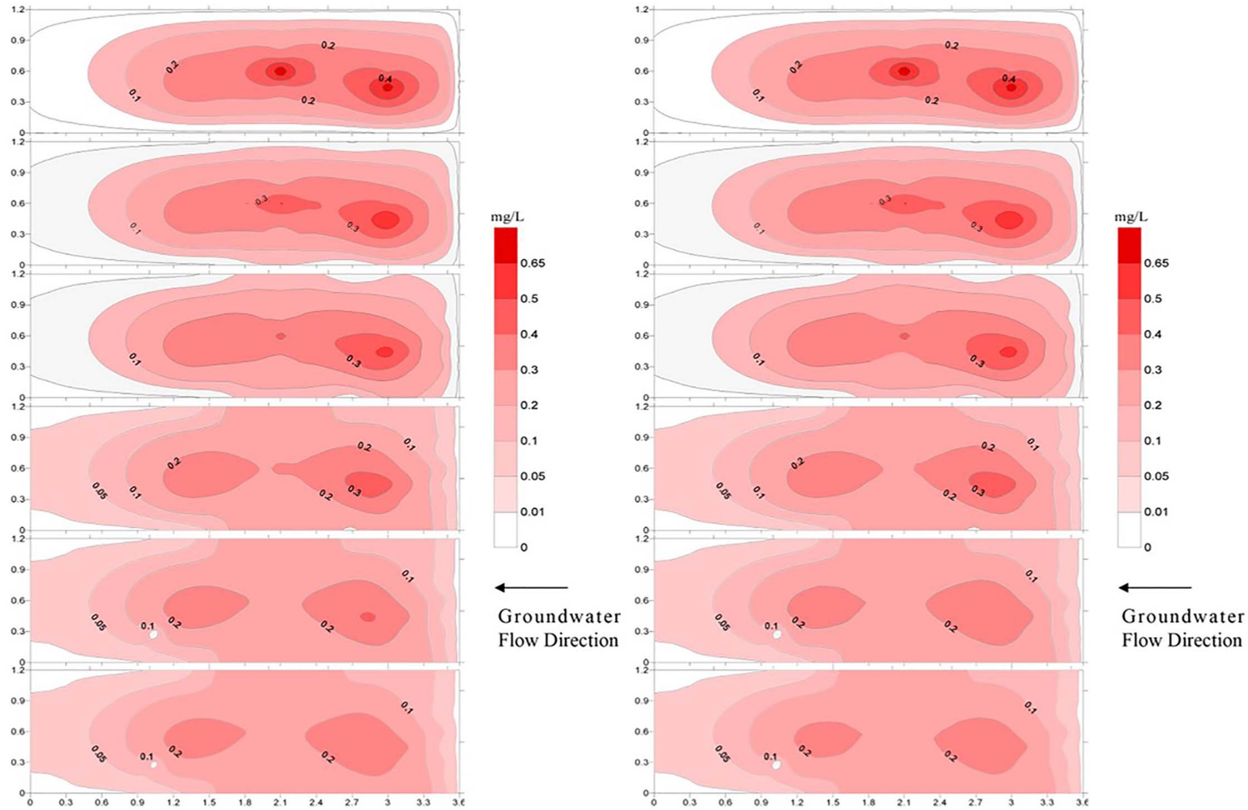


Figure 4. Benzene concentrations on Days 2, 6, 10, 14, 20 and 22 of the DPC system(a) and SADPC system(b)

3.3 Comparison analysis

To examine the remediation efficiency by the DPC system and the SADPC system, nine hypothetical wells (HWs) are selected from the simulation domain and the locations of these HWs can be seen in Figure S4. Figure 5 presents the benzene concentrations of the SADPC system at the nine HWs from Day 2 to Day 22 (Benzene concentrations of the DPC system is shown in Figure S5). In the DPC system, the analysis of the predicted data of benzene concentrations at the nine HWs indicates that the benzene concentrations during the first ten days of remediation decrease slowly or even increase at some locations, and the pumping/injecting rates has been increased to a certain degree accordingly. The signal of an increasing concentration of contaminants triggers the necessary adjustment of the operation. The predicted data also indicates decreases in the contaminant concentrations at almost all locations after ten days of operation, especially at HW-56, HW-102 and HW-106, which show sharp decreases on Day 10 or Day 14. The distances between HW-102/HW-106 and the contaminant source are the same, which can be the reason why these two sites show the similar trend during contaminant degradation. After second-stage rolling optimization, it can be seen that the benzene concentrations at the nine HWs decrease in a faster rate in the SADPC system since Day 12, and reach balances 2 days faster than those in the DPC system. The pumping/injecting rates has been decreased from Day 12 according to the instructions from both systems. Based on the entire degradation process, the benzene concentrations in HW-40, HW-42, HW-48 and HW-52 decline gently, which may due to the fact that these wells are relatively far from the contaminant source. Therefore, the benzene can be removed much easier than those sites that are located near the contaminant source.

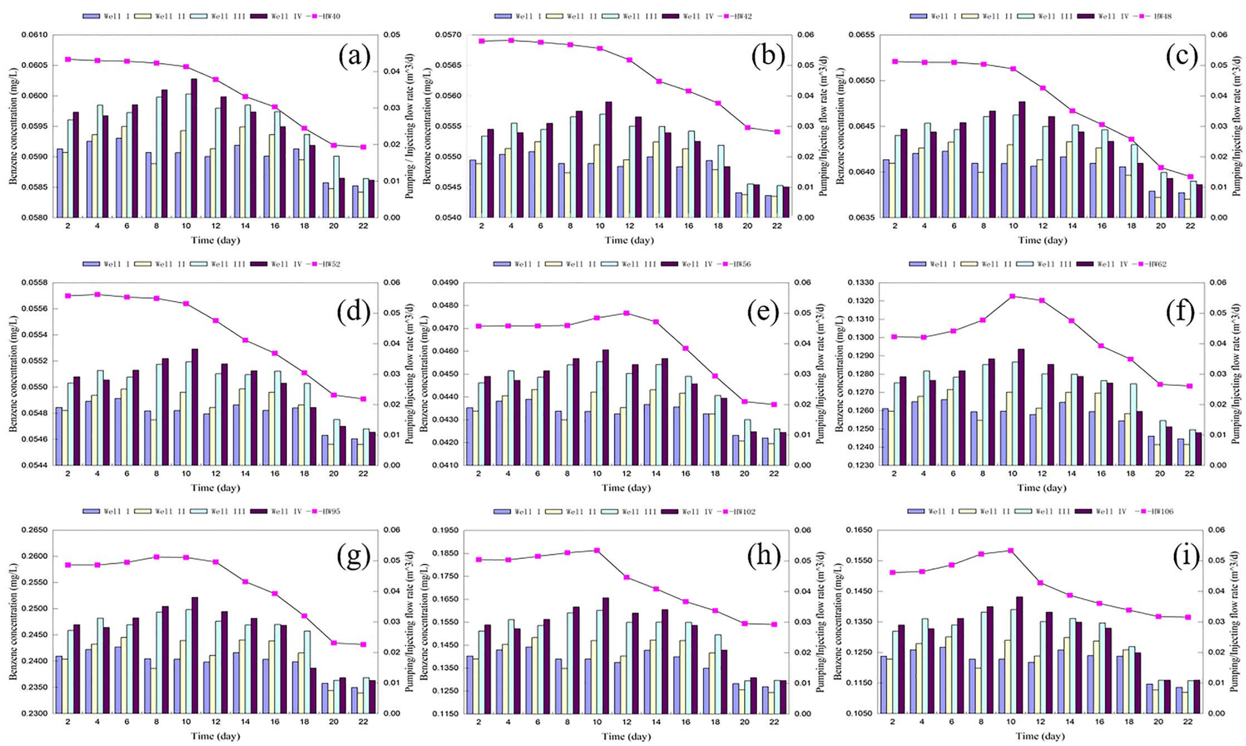


Figure 5. Benzene concentrations of the SADPC system from Day 2 to Day 22, where Figs. (a) to (i) represents the concentrations at HW-40, HW-42, HW-48, HW-52, HW-56, HW-62, HW-95, HW-102, and HW-106

The values of the percentage of benzene mass removal during the remediation process in the SADPC system can be seen in Figure 6. To compare the degradation effect in a straight way, benzene mass removal situations in the SI emulator, FCI optimizer and DPC system were also simulated. Each subgraph constructs an exponentially fitting curve with a 95% confidence band. The coefficient of determination values (R^2) show that all of the degrees of fitting of these models are satisfactory. The removal rate in the SI emulator is the minimum at the beginning of the remediation process, and then the percentage of benzene mass removal rises slowly with fluctuations until the remediation ends on Day 28. The FCI optimizer deals with benzene in groundwater at a relatively steady rate especially from Day 2 to Day 10, which is different from other three methods. Besides, it reaches the equilibrium stage with the removal rate around 82% on Day 26, only faster than the SI emulator. For the DPC system, it is found there is a "remediation plateau period" during the whole process (Day 12 to Day 16), while the percentage of benzene mass removal keeps growing and reaches the cleanup goal on Day 22. About 93% of the benzene mass has been removed by using the DPC system, which means the highest remaining benzene concentration anywhere in the simulation domain has been reduced to below 300 μg/L. Because the benzene removal rates of DPC system cannot meet the expectation from Day 11, an ideal setpoint curve is produced based on the data of the first six points in the DPC system. It should be noted that the SADPC method is applied from Day 14 to Day 28, and the setpoint curve of the SADPC system in Figure 6 is only used for the optimization of the seventh point. The SADPC system improves the biodegradation performance in the DPC system by using the rolling optimization model, which increases the removal rate during the "remediation plateau period" and arrives the cleanup goal on Day 20.

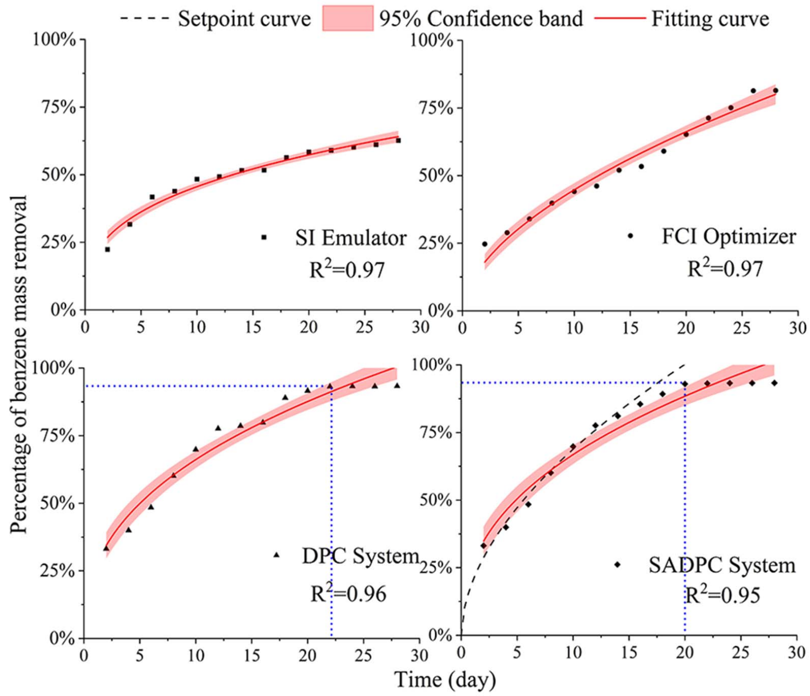


Figure 6. Comparison of predicted benzene removal efficiency by four control systems

The SI emulator only focuses on contamination degradation without considering the cost during the degradation process. For example, the SI emulator improves the pumping/injecting rate when the benzene concentration within the high level, so the removal rate is relatively high at the initial stage. However, the pumping/injecting rate can be reduced as the contaminant concentration deceases. Only 62.6% of the contaminant in the groundwater system has been removed by using the SI emulator which shows that it can hardly remove the benzene effectively. Besides, the whole degradation process in the SI emulator continues for 28 days, which imposes a heavy economical burden on actual stakeholders when applied to practical projects in the future. The optimization objective of the FCI optimizer mainly concentrates on the cost of the operation, which restricts the performance of the benzene removal. Although DPC system improves the degradation rate to some degree, it cannot always meet the expectations of people. In general, the remediation effect of the SADPC system is regarded as the best on the basis of the DPC system, because it fulfils the removal goal within a relatively short period and takes both cost and efficiency into consideration at the same time.

3.4 Policy implication

Bioremediation has demonstrated to be one of the most cost-effective technologies in organic-contaminated groundwater remediation. It can also be combined with other in-situ (or ex-situ) physical or chemical technologies to enhance remediation efficiency, shorten remediation duration, and reduce remediation cost. Historically, numerous studies have been undertaken by concentrating on advancement of new remediation technologies (e.g., adopting highly-efficient microbes and agents, designing new process flows, or optimizing operating conditions) just for improving remediation performance. Unfortunately, few of them attempted to use dynamic process approaches from the perspective of whole process to address the challenge. This study has suggested that a well-designed DPC system could be an easy-to-implement and strong-to-generalize approach compared to those conventional efforts. While this study focused on bioremediation process of benzene-contaminated groundwater, the developed SADPC system can be conveniently extended to many other cases no matter what one will need to challenge: organic or inorganic, physicochemical or biological, and water or soil. The major effort that needs to be accomplished is to construct a set of equations capable of capturing the relationships (frequently called proxy equations) between operating conditions and remediation performance. Nonetheless, as this study is a first attempt, much improvement will be desired, for example, simplifying the system framework to alleviate computational effort, strengthening the error information feedback to shorten decision duration, and introducing stochastic analysis to further mitigate the uncertainty impact.

AI plays an important role in running this SADPC system, which includes stepwise inference, stepwise filtered-clustering inference, genetic algorithm, etc. As conventional physically based models can hardly be directly used by SADPC considering the independence of prior assumptions for model forms (He et al., 2008), these machinery-learning similar inference methods are introduced to create a set of computation-fast and accuracy-reliable proxy equations to replace the conventional physical model. Note that a physical model is not a must by all the cases (this study uses the physical model aiming to generate a substantial number of statistical samples to obtain proxy equations). This implies that AI techniques have high potential to be used in remediation studies and practices because of their strong capabilities of convenient modeling (particularly in modeling highly nonlinear input-output relations), automatic learning, self-adaptation, and reliable generalization. It is desired that more state-of-the-art AI techniques be introduced in the SADPC system particularly including those intelligent denoising, error correction, and decision-making. This will much increase the performance of dynamic process control and enrich available control approaches.

An obvious knowledge gap implied in this study is the lack of related guidelines when designing a SADPC system. Historically, various guidelines on groundwater and soil remediation have been proposed in these past years at the national, provincial (or state-), and municipal levels. In terms of these guidelines, one can easily know what technologies can be used, how the remediation wells should be configured, which criteria should be satisfied after remediation, etc. Without the guidelines, there will be a difficulty in guaranteeing the stability, maturity, and reliability of a newly designed SADPC system, probably leading to extraordinary carefulness or even refuse of the potential users. This suggested that guidelines on remediation-oriented SADPC systems could be tentatively investigated, developed, and applied in future effort. This work is significant for offering users a set of principles or rules to follow when designing a SADPC system and helping them clarify the problems such as what procedures should be implemented when designing and running a SADPC system, how the control errors should be guaranteed during the whole remediation process, and what criteria could be adopted to evaluate the control performance.

4 Conclusions

A system was developed to improve the removal rate of benzene for enhancing in-situ bioremediation, which can effectively reflect the dynamics and complexity of the biodegradation process and realize the control for the remediation system based on the feedback information. The insights from this study can suggest to decision makers that guidelines and policies on remediation-oriented SADPC systems could be tentatively investigated, developed, and applied in future effort.

- Results from the error analysis of the contaminant fate and transport model show that the model agrees well with the data obtained from the pilot experiments, so it can be used for developing the SADPC system.
- The SADPC system is consist of an FCI simulator, a SI emulator, an MPC controller, a 1st-stage optimizer, and a 2nd-stage optimizer.
- The SADPC system improves the biodegradation performance in the DPC system by using the rolling optimization model, which increases the removal rate during the "remediation plateau period" and arrives at the cleanup goal earlier than the DPC system, as well as the SI emulator and the FCI optimizer.

Acknowledgments

The authors acknowledge funding from National Key Research and Development Program of China (Grant No. 2020YFC1807904).

References

Ahmed, W., & Rodriguez, J. (2020). A model predictive optimal control system for the practical automatic start-up of anaerobic digesters. *Water Research*, 174. 115599. https://doi.org/10.1016/j.watres.2020.115599.

Akyol, K. (2020). Comparing of deep neural networks and extreme learning machines based on growing and pruning approach. *Expert Systems with Application*, 140(Feb.),112875.1-112875.7. https://doi.org/10.1016/j.eswa.2019.112875.

Albers, C. N., Feld, L., Ellegaard-Jensen, L., & Aamand, J. (2015). Degradation of trace concentrations of the persistent groundwater pollutant 2,6-dichlorobenzamide (BAM) in bioaugmented rapid sand filters. *Water Research*, 83, 61-70. https://doi.org/10.1016/j.watres.2015.06.023.

Bashivan, P., Kar, K., & DiCarlo, J. J. (2019). Neural population control via deep image synthesis. *Science*, 364(6439), 453-+. https://doi.org/10.1126/science.aav9436.

Bear, J. (1979). *Hydraulics of Ground Water*. Newyork: Dover Publications.

Bechet, Q., Shilton, A., & Guieysse, B. (2016), Maximizing Productivity and Reducing Environmental Impacts of Full-Scale Algal Production through Optimization of Open Pond Depth and Hydraulic Retention Time. *Environmental Science Technology*, 50(7), 4102-4110. https://doi.org/10.1021/acs.est.5b05412.

Chang, H. L., & Alvarez-Cohen, L. (1995). Model for the cometabolic biodegradation of chlorinated organics. *Environmental Science & Technology*, 29(9), 2357. https://doi.org/10.1021/es00009a031.

Chang, H. L., & Alvarez-Cohen, L. (2010). Transformation capacities of chlorinated organics by mixed cultures enriched on methane, propane, toluene, or phenol. *Biotechnology & Bioengineering*, 45(5), 440-449. https://doi.org/10.1002/bit.260450509.

Chiandussi, G., Codegone, M., Ferrero, S., & Varesio, F. E. (2012). Comparison of multi-objective optimization methodologies for engineering applications. Computers & Mathematics with *Applications*, 63(5), 912-942. https://doi.org/10.1016/j.camwa.2011.11.057.

de Blanc, P. C. (1998). Development and demonstration of a biodegradation model for non-aqueous-phase liquids in groundwater, (Doctoral dissertation). Austin, Texas: The University of Texas at Austin.

Delshad, M., Pope, G.A., & Sepehrnoori, K. (1996). A compositional simulator for modeling surfactant enhanced aquifer remediation, 1 formulation. *Journal of Contaminant Hydrology*, 23(4), 303-327. https://doi.org/10.1016/0169-7722(95)00106-9.

Diangelakis, N. A., Avraamidou, S., & Pistikopoulos, E. N. (2016). Decentralized Multiparametric Model Predictive Control for Domestic Combined Heat and Power Systems. *Industrial & Engineering Chemistry Research*, 55(12), 3313-3326. https://doi.org/10.1021/acs.iecr.5b03335.

Faust, C. R., Guswa, J. H., & Mercer, J. W. (1989). Simulation of three-dimensional flow of immiscible fluids within and below the unsaturated zone. *Water Resources Research*, 25(12), 2449-2464. https://doi.org/10.1029/WR025i012p02449.

Gerber, S., & Horenko, I. (2015). Improving clustering by imposing network information. *Science Advances*, 1(7). https://doi.org/10.1126/sciadv.1500163.

Gopalakrishnan, G., Minsker, B. S., & Valocchi, A. J. (2011). Monitoring Network Design for Phytoremediation Systems Using Primary and Secondary Data Sources. *Environmental Science & Technology*, 45(11), 4846-4853. https://doi.org/10.1021/es1042657.

Gorelick, S. M., & Zheng, C. M. (2015). Global change and the groundwater management challenge. *Water Resources Research*, 51(5), 3031-3051. https://doi.org/10.1002/2014WR016825.

Greenland, S., Daniel, R., & Pearce, N. (2016). Outcome modelling strategies in epidemiology: traditional methods and basic alternatives. *International Journal of Epidemiology*, 45(2), 565-575. https://doi.org/10.1093/ije/dyw040.

Hao, P. Y., & Chiang, J. H. (2006). Pruning and model-selecting algorithms in the RBF frameworks constructed by support vector learning. *International Journal of Neural Systems*, 16(4), 283-293. https://doi.org/10.1142/S0129065706000688.

He, L., 2008. Development of integrated simulation and optimization models for petroleum-contaminated groundwater remediation management under various uncertainties, (Doctoral dissertation). Regina, Saskatchewan: The University of Regina.

He, L., Huang, G. H., & Lu, H. W. (2008a). Health-Risk-Based Groundwater Remediation System Optimization through Clusterwise Linear Regression. *Environmental Science & Technology*, 42(24), 9237-9243. https://doi.org/10.1021/es800834x.

He, L., Huang, G. H., Lu, H. W., & Zeng, G.M. (2008b). Optimization of surfactant-enhanced aquifer remediation for a laboratory BTEX system under parameter uncertainty.

Environmental Science & Technology, 42(6), 2009-2014. https://doi.org/10.1021/es071106y.
He, L., Huang, G. H., Zeng, G. M., Lu, & H. W. (2008c). An integrated simulation, inference, and optimization method for identifying groundwater remediation strategies at petroleum-contaminated aquifers in western Canada. *Water Research*, 42(10-11), 2629-2639. https://doi.org/10.1016/j.watres.2008.01.012.
Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems.* Ann Arbor: University of Michigan Press.
Hou, Z., Chi, R., & Gao, H. (2017). An Overview of Dynamic-Linearization-Based Data-Driven Control and Applications. *Ieee Transactions on Industrial Electronics*, 64(5), 4076-4090. https://doi.org/10.1109/TIE.2016.2636126.
Hu, Y., Patmonoaji, A., Xu, H., Kaito, K., & Suekane, T. (2021). Pore-scale investigation on nonaqueous phase liquid dissolution and mass transfer in 2d and 3d porous media. *International Journal of Heat and Mass Transfer*, 169, 120901. https://doi.org/10.1016/j.ijheatmasstransfer.2021.120901.
Huang, G. H., Huang, Y. F., Wang, G. Q., & Xiao, H.N. (2006). Development of a forecasting system for supporting remediation design and process control based on NAPL-biodegradation simulation and stepwise-cluster analysis. *Water Resources Research*, 42(6). https://doi.org/10.1029/2005WR004006.
Huang, Y. F., Wang, G. Q., Huang, G. H., Xiao, H. N., & Chakma, A. (2008). IPCS: An integrated process control system for enhanced in-situ bioremediation. *Environmental pollution*, 151(3), 460-469. https://doi.org/10.1016/j.envpol.2007.04.010.
Itoh, Y., & Adachi, M. (2017). *Reconstruction of bifurcation diagrams using an extreme learning machine with a pruning algorithm*, Paper presented at International Joint Conference on Neural Networks (IJCNN), Institute of Electrical and Electronics Engineers Inc., Anchorage, AK, United states.
Jacobs, O. L. R. (1993). *Introduction to Control Theory*. Oxford: Oxford University Press,.
Jimenez, N., Vinas, M., Sabate, J., Diez, S., & Bayona, J. M. (2006). The prestige oil spill. 2. enhanced biodegradation of a heavy fuel oil under field conditions by the use of an oleophilic fertilizer. *Environmental Science & Technology*, 40(8), 2578-2585. https://doi.org/10.1021/es052370z.
Kuo, J. T., Wang, Y. Y., & Lung, W. S. (2006). A hybrid neural–genetic algorithm for reservoir water quality management. *Water research*, 40(7), 1367-1376. https://doi.org/10.1016/j.watres.2006.01.046.
Lenczewski, M., Jardine, P., McKay, L., & Layton, A. (2003). Natural attenuation of trichloroethylene in fractured shale bedrock. *Journal of Contaminant Hydrology*, 64(3-4), 151-168. https://doi.org/10.1016/S0169-7722(02)00090-6.
Li, Y. S., Abriola, L. M., Phelan, T. J., Ramsburg, C.A., & Pennell, K. D. (2007). Experimental and numerical validation of the total trapping number for prediction of DNAPL mobilization. *Environmental Science & Technology*, 41(23), 8135-8141. https://doi.org/10.1021/es070834i.
Liang X., Devine C. E., Nelson J., Lollar, B. S., Zinder, S., & Edwards, E. A. (2013). Anaerobic conversion of chlorobenzene and benzene to CH4 and CO2 in bioaugmented microcosms. *Environmental Science & Technology*, 47(5): 2378-2385. https://doi.org/10.1021/es3043092.

Liao, R., Jin, Z., Chen, M., & Li, S. (2020). An integrated approach for enhancing the overall performance of constructed wetlands in urban areas. *Water Research*, 187, 116443. https://doi.org/10.1016/j.watres.2020.116443.

Liu, L., Zhou, Q., Liang, H. J., & Wang, L. J. (2017). Stability and Stabilization of Nonlinear Switched Systems Under Average Dwell Time. *Applied Mathematics and Computation*, 298, 77-94. https://doi.org/10.1016/j.amc.2016.11.006.

Liu, Y. Q., Ganigue, R., Sharma, K., & Yuan, Z. G. (2016). Event-driven model predictive control of sewage pumping stations for sulfide mitigation in sewer networks. *Water Research*, 98, 376-383. https://doi.org/10.1016/j.watres.2016.04.039.

Matott, L. S., Bartelt-Hunt, S. L., Rabideau, A. J., & Fowler, K. (2006). Application of heuristic optimization techniques and algorithm tuning to multilayered sorptive barrier design. *Environmental Science & Technology*, 40(20), 6354-6360. https://doi.org/10.1021/es052560+.

Maybeck, P. S. (1979). *Stochastic models, estimation and control.* (Vol. 1). London: Academic Press, INC.

Mayne, D. Q. (2014). Model predictive control: Recent developments and future promise. *Automatica* 50(12), 2967-2986. https://doi.org/10.1016/j.automatica.2014.10.128.

McDonald, M. G., & Harbaugh, A. W. (1988). *A modular three-dimensional finite-difference groundwater flow model*. Washington, DC: United States Government Printing Office (USGPO).

Meray, A. O., Sturla, S., Siddiquee, M. R., Serata, R., Uhlemann, S., Gonzalez-Raymat, H., Denham, M., Upadhyay, H., Lagos, L. E., Eddy-Dilek, C., & Wainwright, H. M. (2022). PyLEnM: A Machine Learning Framework for Long-Term Groundwater Contamination Monitoring Strategies. *Environmental Science & Technology*, 56(9), 5973-5983. https://doi.org/10.1021/acs.est.1c07440.

Miller, T. H., Baz-Lomba, J. A., Harman, C., Reid, M. J., Owen, S. F., Bury, N. R., Thomas, K. V., & Barron, L. P. (2016). The First Attempt at Non-Linear in Silico Prediction of Sampling Rates for Polar Organic Chemical Integrative Samplers (POCIS). *Environmental Science & Technology*, 50(15), 7973-7981. https://doi.org/10.1021/acs.est.6b01407.

Niswonger, R. G., & Prudic, D. E. (2013). *Modeling variably saturated flow using kinematic waves in MODFLOW*. American Geophysical Union.

Opher, T., & Ostfeld, A. (2011). A coupled model tree (MT) genetic algorithm (GA) scheme for biofouling assessment in pipelines. *Water research*, 45(18), 6277-6288. https://doi.org/10.1016/j.watres.2011.09.037.

Passino, K. M. (2002). Biomimicry of bacterial foraging for distributed optimization and control. *Ieee Control Systems Magazine* 22(3), 52-67. https://doi.org/10.1109/mcs.2002.1004010.

Pizzagalli, D. U., Gonzalez, S. F., & Krause, R. (2019). A trainable clustering algorithm based on shortest paths from density peaks. *Science Advances*, 5(10). https://doi.org/10.1126/sciadv.aax3770.

Rao, C. R. (1952). Advanced statistical methods in biometric research. *Biometrics*, 28(1), 253. https://doi.org/10.2307/2528984.

Rittmann, B. E., Henry, B., Odencrantz, J. E., & Sutfin, J. A. (1991). Biological fate of a polydisperse acrylate polymer in anaerobic sand-medium transport. *Biodegradation*, 2(3), 171-179. https://doi.org/10.1007/BF00124491.

Schaerlaekens, J., Carmeliet, J., & Feyen, J. (2005). Multi-objective optimization of the setup of a surfactant-enhanced DNAPL remediation. *Environmental Science & Technology*, 39(7), 2327-2333. https://doi.org/10.1021/es049148z.

Shen, H., Li, F., Wu, Z. G., Park, J. H., & Sreeram, V. (2018). Fuzzy-Model-Based Nonfragile Control for Nonlinear Singularly Perturbed Systems With Semi-Markov Jump Parameters. *Ieee Transactions on Fuzzy Systems*, 26(6), 3428-3439. https://doi.org/10.1109/TFUZZ.2018.2832614.

Siade, A. J., Cui, T., Karelse, R. N., & Hampton, C. (2020). Reduced-Dimensional Gaussian Process Machine Learning for Groundwater Allocation Planning Using Swarm Theory. *Water Resources Research*, 56(3). https://doi.org/10.1029/2019wr026061.

Stentoft, P. A., Munk-Nielsen, T., Moller, J. K., Madsen, H., Valverde-Perez, B., Mikkelsen, P. S., & Vezzaro, L. (2021). Prioritize effluent quality, operational costs or global warming? - Using predictive control of wastewater aeration for flexible management of objectives in WRRFs. *Water Research*, 196. https://doi.org/10.1016/j.watres.2021.116960.

Stramer, Y., Brenner, A., Cohen, S. B., & Oron, G. (2010). Selection of a Multi-Stage System for Biosolids Management Applying Genetic Algorithm. *Environmental Science & Technology*, 44(14), 5503-5508. https://doi.org/10.1021/es902981t.

Sun, J., Miao, Z., Gong, D. W., Zeng, X. J., Li, J. Q., & Wang, G. G. (2020). Interval Multiobjective Optimization With Memetic Algorithms. *Ieee Transactions on Cybernetics*, 50(8), 3444-3457. https://doi.org/10.1109/tcyb.2019.2908485.

Sun, K., Yu, Y., & Huang, Z. (2015). *A Generalized Pruning Algorithm for Extreme Learning Machine*. Paper presented at IEEE International Conference on Information and Automation 2015, Lijiang, PEOPLES R CHINA.

Umar, M. F., Rafatullah, M., Abbas, S. Z., Ibrahim, M., & Ismail, N. (2021). Enhanced benzene bioremediation and power generation by double chamber benthic microbial fuel cells fed with sugarcane waste as a substrate. *Journal of Cleaner Production*, 310(4), 127583. https://doi.org/10.1016/j.jclepro.2021.127583.

Wang, W. J., Tian, G. D., Chen, M. N., Tao, F., & Zhang, C. Y., Ai-Ahmari, A., Li, Z. W., & Jiang, Z. G. (2020). Dual-objective program and improved artificial bee colony for the optimization of energy-conscious milling parameters subject to multiple constraints. *Journal of Cleaner Production*, 245. https://doi.org/10.1016/j.jclepro.2019.118714.

Wolicka, D., Suszek, A., Borkowski, A., & Bielecka, A. (2009). Application of aerobic microorganisms in bioremediation in situ of soil contaminated by petroleum products. *Bioresource Technology*, 99(13), 3221-3227. https://doi.org/10.1016/j.biortech.2009.02.020.

Xin, B. P., Wu, C. H., Wu, C. H., & Lin, C. W. (2013). Bioaugmented remediation of high concentration btex-contaminated groundwater by permeable reactive barrier with immobilized bead. *Journal of Hazardous Materials*, 244-245(JAN.15), 765-772. https://doi.org/10.1016/j.jhazmat.2012.11.007.

Yang, C., Liu, S., Su, Y., Chen, Y., Lin, C., & Lin, K. (2019). Bioremediation capability evaluation of benzene and sulfolane contaminated groundwater: determination of bioremediation parameters. *The Science of the Total Environment*, 648, 811-818. https://doi.org/10.1016/j.scitotenv.2018.08.208.

Zeng, J., & Liu, J. F. (2015). Economic Model Predictive Control of Wastewater Treatment Processes. *Industrial & Engineering Chemistry Research*, 54(21), 5710-5721. https://doi.org/10.1021/ie504995n.

Zhan, K., Zhang, C. Q., Guan, J. P., & Wang, J. S. (2018). Graph Learning for Multiview Clustering. *Ieee Transactions on Cybernetics*, 48(10), 2887-2895. https://doi.org/10.1109/tcyb.2017.2751646.

Zhang, C., Wu, D. J., & Ren, H. X. (2020). Bioremediation of oil contaminated soil using agricultural wastes via microbial consortium. *Scientific Reports*, 10(1). https://doi.org/10.1038/s41598-020-66169-5.

Zhang, C., Yoon, H., Werth, C. J., Valocchi, A. J., Basu, N. B., & Jawitz, J. W. (2008). Evaluation of simplified mass transfer models to simulate the impacts of source zone architecture on nonaqueous phase liquid dissolution in heterogeneous porous media. *Journal of Contaminant Hydrology*, 102(1-2), 49-60. https://doi.org/10.1016/j.jconhyd.2008.05.007.

Zou, Y., Huang, G. H., & Nie, X. H. (2009). Filtered Stepwise Clustering Method for Predicting Fate of Contaminants in Groundwater Remediation Systems: A Case Study in Western Canada. *Water Air and Soil Pollution*, 199(1-4), 389-405. https://doi.org/0.1007/s11270-008-9887-5.

Figure1.

Optimization Loop
FCI Simulator
$X_p(t+1)$
1st-Stage Optimizer
$U'(t)$
$U(t)$
$X(t)$
$U^{①*}(t)$
$U^{①*}(t)$
SI Emulator
MPC Controller
$U^{②*}(t)$
Biodegradation Process
$X^*(t+1)$
$U^{①*}(t)$
FCI Simulator
$X(t)$
$X_p(t+1)$
$U'(t)$
Optimization Loop
-
Setpoint Curve
$X_r(t+1)$
+
$e(t+1)$
2nd-Stage Optimizer
$U^{②*}(t)$

(a) (b)

Figure2.

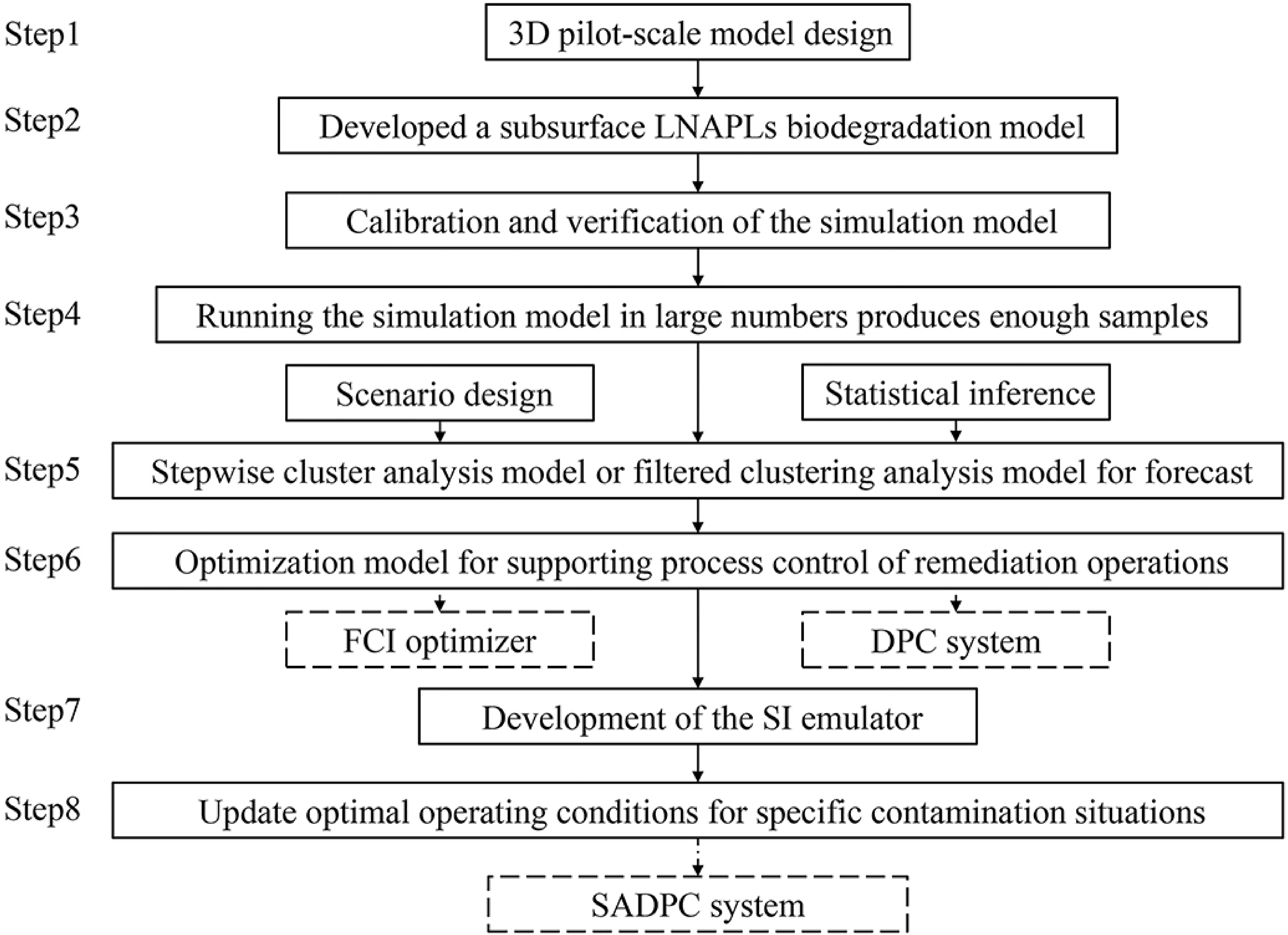

Step1
3D pilot-scale model design
Step2
Developed a subsurface LNAPLs biodegradation model
Step3
Calibration and verification of the simulation model
Step4
Running the simulation model in large numbers produces enough samples
Scenario design
Statistical inference
Step5
Stepwise cluster analysis model or filtered clustering analysis model for forecast
Step6
Optimization model for supporting process control of remediation operations
FCI optimizer
DPC system
Step7
Development of the SI emulator
Step8
Update optimal operating conditions for specific contamination situations
SADPC system

Figure3.

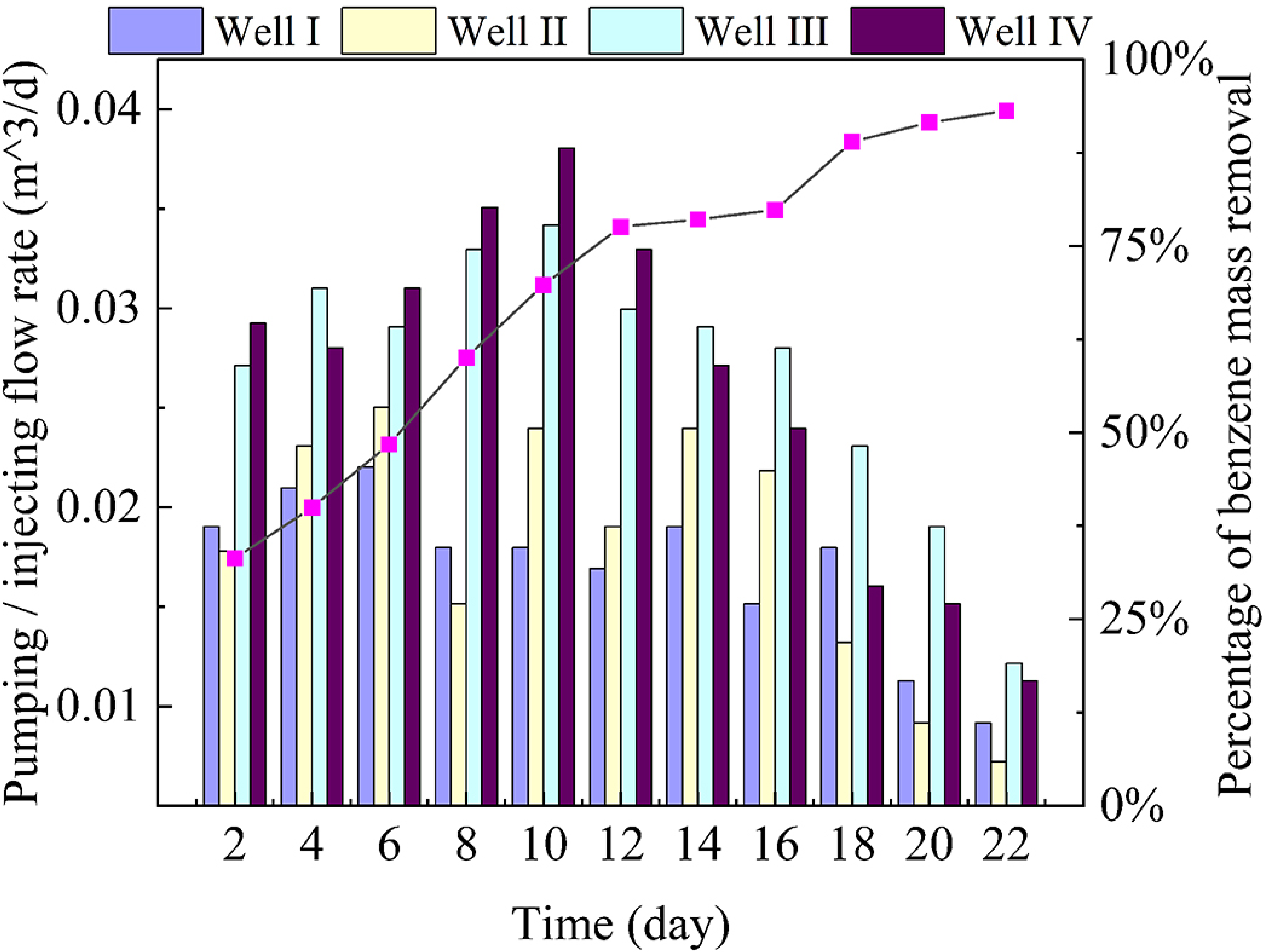
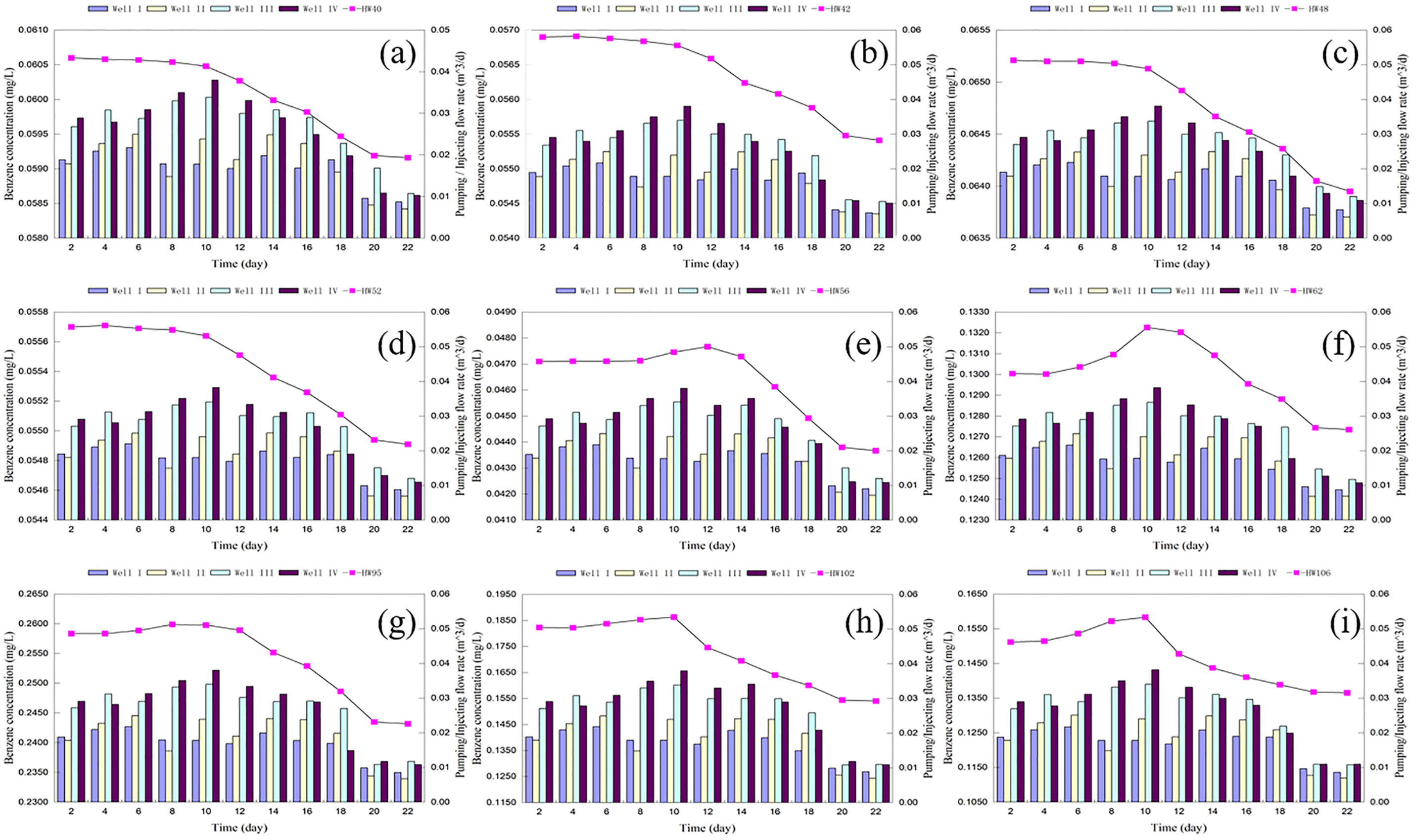

Well I
Well II
Well III
Well IV
Pumping / injecting flow rate (m^3/d)
0.01
0.02
0.03
0.04
Percentage of benzene mass removal
0%
25%
50%
75%
100%
2
4
6
8
10
12
14
16
18
20
22
Time (day)

Figure4.

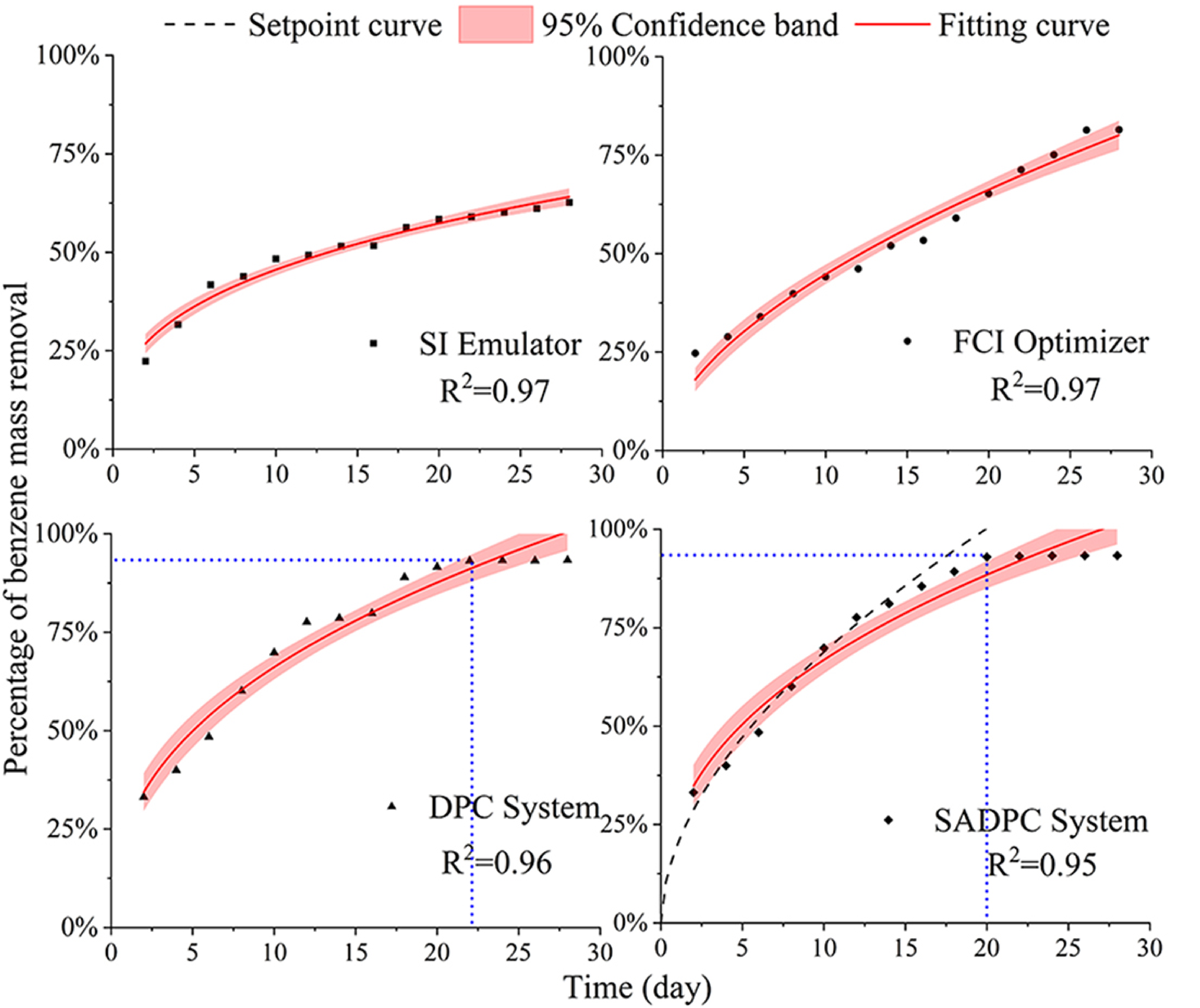
mg/L
0.65
0.5
0.4
0.3
0.2
0.1
0.05
0.01
0
Groundwater Flow Direction
0 0.3 0.6 0.9 1.2 1.5 1.8 2.1 2.4 2.7 3 3.3 3.6
0 0.3 0.6 0.9 1.2

Figure5.



Well I
Well II
Well III
Well IV
HW40
HW42
HW48
HW52
HW56
HW62
HW95
HW102
HW106
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
Benzene concentration (mg/L)
Pumping / Injecting flow rate (m^3/d)
Pumping/Injecting flow rate (m^3/d)
Time (day)

Figure6.



Setpoint curve
95% Confidence band
Fitting curve
Percentage of benzene mass removal
SI Emulator
R^2=0.97
FCI Optimizer
R^2=0.97
DPC System
R^2=0.96
SADPC System
R^2=0.95
Time (day)
100%
75%
50%
25%
0%
0
5
10
15
20
25
30