https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

A data-driven method coupling multiple physical constraints for correcting structural errors in groundwater contaminant transport models

Jinglong Tian[1], Xiankui Zeng[1*], Yue Pan[2], Dong Wang[1], and Jichun Wu[1]

[1] School of Earth Sciences and Engineering, Key Laboratory of Surficial Geochemistry, Ministry of Education, Nanjing University, Nanjing, China

[2] College of Transport Science and Engineering, Nanjing Tech University, Nanjing, China

Corresponding author: Xiankui Zeng (xiankuizeng@nju.edu.cn)

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

Abstract

Model structural errors are pervasive in groundwater contaminant transport modeling under complex environmental conditions, hindering accurate prediction of contamination transport. Data-driven methods (DDMs) coupled with physical constraints provide an effective approach for correcting structural errors and improving prediction. However, in multicomponent reactive transport systems, multiple physical mechanisms must be satisfied simultaneously, whereas existing DDMs have limited capacity to effectively couple multiple physical constraints. To address this challenge, this study proposes a general method for correcting structural errors in groundwater models. A combined likelihood function is constructed and sub-likelihood weights are dynamically updated to effectively couple multiple physical constraints. The method is evaluated using a synthetic three-dimensional tetrachloroethylene reactive transport simulation and a cadmium-phosphate cotransport sand column experiment. These tests systematically assess the effects of coupling single versus multiple physical constraints on structural error correction and predictive performance. The results show that coupling multiple constraints can constrain parameter identification, reduce predictive uncertainty, and more comprehensively improve model predictions. Appropriate physical constraints function analogously to incorporating additional observations. Moreover, coupling multiple physical constraints results in a simpler form of structural error in the calibrated groundwater model, making it easier to characterize, thereby enhancing prediction accuracy and physical consistency. The proposed dynamic updating and stopping criterion of sub-likelihood weights maintains a balance between multiple physical constraints and observations, improving the robustness of parameter identification and constraint enforcement. Overall, the proposed DDM coupled with multiple physical constraints provides a general framework for correcting structural errors in complex groundwater contaminant transport models.

Keywords: Gaussian process regression; Model structural error; Multiple physical constraints; Mass conservation

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

1. Introduction

Numerical groundwater models are essential tools for predicting contaminant transport processes in groundwater and play a critical role in risk assessment and remediation decision-making at contaminated sites (Kim et al., 2026; Wen et al., 2025). With increasing demands for efficient and sustainable remediation of groundwater contamination, improving the predictive accuracy of complex migration and transformation processes of contaminants in aquifers, together with conducting reliable uncertainty analysis, has become imperative (Zeeshan and Ruhl, 2023). Although extensive studies have focused on parameter uncertainty (Hendrikx et al., 2026; Engdahl, 2025), comparatively limited attention has been devoted to model structural error. Owing to the complexity of groundwater systems and limitations in human understanding, constructed groundwater models are often simplified or even misspecified representations of real systems, resulting in the widespread presence of structural error (Refsgaard et al., 2006; Neuman and Wierenga, 2003). Such structural deficiencies lead to systematic errors in model predictions (Watson et al., 2013; Ye et al., 2004), thereby reducing the reliability of prediction-informed decisions (Sun et al., 2021; Gupta et al., 2012).

In recent years, data-driven methods (DDMs) have been widely employed to correct model structural errors and enhance the predictive performance of physics-based models (Yu et al., 2022; Hu et al., 2021; Kasiviswanathan et al., 2019; Pathiraja et al., 2018). These approaches construct an error-correcting complementary model to explicitly account for systematic discrepancies in model predictions, such that the corrected predictions consist of the physical model, the structural error model, and the measurement error (Xu et al., 2017; Xu and Valocchi, 2015). Previous studies have demonstrated that applying DDMs to correct structural error can substantially improve the predictive accuracy of groundwater head, soil moisture, and contaminant concentration (Pan et al., 2023; Sun et al., 2021; Zhang et al., 2019). However, the error-correction models constructed by these approaches remain essentially black-box statistical models that lack necessary physical constraints. Consequently, they may

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

yield predictions that violate fundamental physical principles, such as contaminant concentrations exceeding physically plausible ranges or violating mass conservation, thereby weakening the plausibility and reliability of the predictions.

Incorporating physical constraints into DDMs provides an effective means of enhancing the physical consistency of predictions. Previous studies have parameterized physical constraints and formulated them as constraint conditions in optimization problems to achieve coupling of physical mechanisms (Ayensa-Jiménez et al., 2018; He and Xiu, 2016). However, when multiple physical constraints are considered simultaneously, the number of unknown parameters may exceed the number of physical constraints, thereby increasing predictive uncertainty. Wu et al. (2024) incorporated the energy conservation constraint into parameter identification by jointly considering the accumulated violation of the energy conservation constraint and prediction error. Nevertheless, under multiple physical constraints, different physical constraints may compete with one another, making it difficult for predictions to simultaneously satisfy all physical mechanisms. In the groundwater field, Xu and Valocchi (2015) coupled model predictions with the water balance equation through recalibration. However, imposing multiple physical constraints through recalibration may drive the physical model parameters to overfit in order to satisfy all constraints and result in numerical instability or infeasible solutions, thereby limiting the applicability of the approach. Overall, although existing methods can couple a single physical constraint, there remains a lack of a general method capable of simultaneously and effectively coupling multiple physical constraints.

The rapid development of physics-informed machine learning (PIML) offers a potential pathway for coupling multiple physical constraints. PIML typically incorporates physical mechanisms at three stages: during data preparation, by selecting or augmenting training data based on physical constraints (Melching et al., 2023; Zhang et al., 2021); during model construction, by introducing physics-constrained layers (Mendenhall et al., 2024; Saha et al., 2021); and during model training, by designing specific loss functions or inference algorithms to enforce physical constraints (Zhang

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

et al., 2023; Guo et al., 2022; Zhang et al., 2022). To date, PIML has been widely applied in hydrology, engineering, and materials science (Cao and Weng, 2024; Cooper et al., 2023; Lu et al., 2022), demonstrating its potential to integrate multiple physical constraints simultaneously, such as mass balance, monotonicity, and boundary conditions. However, existing PIML approaches still face limitations in correcting model structural errors. This is primarily because PIML typically constrains only the error-correction model itself rather than the entire error-correction process. As a result, the corrected predictions, which consist of the physical model, the structural error model, and measurement error, may still violate fundamental physical principles. Moreover, existing PIML approaches lack effective strategies for selecting appropriate physical constraints tailored to model structural errors, further limiting their potential to enhance predictive reliability and physical consistency.

By constructing a combined likelihood function within a Bayesian uncertainty analysis framework, Tian et al. (2026) successfully coupled Gaussian process regression with a mass conservation constraint to correct structural errors in groundwater models. However, in complex groundwater systems involving multicomponent reactive transport processes, it is often necessary to simultaneously constrain the physical mechanisms associated with multiple components. In such cases, because different physical constraints represent distinct physical meanings, the magnitudes of their corresponding likelihood functions may differ by several orders of magnitude. Simply combining these likelihood functions may cause certain constraints to dominate or be neglected, thereby hindering the effective coupling of multiple physical constraints. Therefore, developing a general structural error correction method capable of coupling multiple physical constraints is essential for improving the accuracy and reliability of transport simulations in complex groundwater contamination systems.

This study proposes a general physics-constrained DDM framework that simultaneously embeds multiple physical constraints into the structural error correction process. The constructed combined likelihood function consists of two components: a

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

conventional likelihood function that quantifies the fit between predictions and observations, and a second component comprising multiple sub-likelihood functions that measure the extent to which predictions satisfy individual physical constraints. In addition, an adaptive weight-updating strategy with a stopping criterion is introduced. During parameter identification, the sub-likelihood function weights are updated according to their relative magnitudes, thereby achieving a reasonable balance between data fitting and multiple physical constraints. The effectiveness of the proposed method is evaluated using two groundwater contaminant transport simulation cases: a synthetic three-dimensional tetrachloroethylene (PCE) reactive transport simulation and a cadmium-phosphate cotransport sand column experiment. The results indicate that the proposed physics-constrained DDM framework effectively couples multiple physical constraints, significantly improving predictive performance of groundwater models and providing a reliable basis for risk assessment and remediation decision-making at complex contaminated sites.

The remainder of this paper is organized as follows. Section 2 presents the Bayesian data-driven framework, the formulation for coupling multiple physical constraints, and the weight-updating strategy with a stopping criterion. Sections 3 and 4 describe the configurations of the two case studies and the implementation of the physics-constrained DDM framework, respectively. Section 5 presents and discusses the results. Section 6 summarizes the main conclusions.

2. Methods

2.1. Bayesian Data-driven Methods

For a groundwater system, the observed data D can be expressed as (Kennedy and O'Hagan, 2001):

$$D = f(\theta) + b(x, \varphi) + \varepsilon \tag{1}$$

where f denotes the groundwater model and θ represents the model parameters. $b(x, \varphi)$ denotes the systematic prediction error described by Gaussian process regression (GPR),

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

where x and φ represent the inputs and hyperparameters of the GPR model, respectively. ε denotes random measurement error.

In this study, Gaussian process regression (GPR) is employed to construct the structural error correction model $b(x,\varphi)$. It is assumed that $b(x,\varphi)$ follows a multivariate Gaussian distribution $N(\mu,C)$, which is fully characterized by the mean function $\mu(x)=E(b(x))$ and the covariance function $k(x,x')=E((b(x)-\mu(x))(b(x')-\mu(x')))$. The mean function μ is set to zero, and a Matérn covariance kernel with smoothness parameter ν=5/2 is adopted (Rasmussen and Williams, 2005):

$$C_{i,j}=k(x_i,x_j)=\sigma^2(1+\frac{\sqrt{5}(x_i-x_j)^T(x_i-x_j)}{\lambda}+\frac{5((x_i-x_j)^T(x_i-x_j))^2}{3\lambda^2}) \times\exp(-\frac{(x_i-x_j)^T(x_i-x_j)}{\lambda^2})+\sigma_\varepsilon^2 I \quad (2)$$

where I denotes an indicator function that equals 1 if $i=j$ and 0 otherwise. The hyperparameter φ consists of λ, σ^2 and σ_ε^2, where λ denotes the characteristic length scale, σ^2 represents the marginal variance of $b(x)$, and σ_ε^2 denotes the variance of measurement error.

The groundwater model parameters θ and the GPR hyperparameters φ are commonly assumed to be independent in the prior distribution. Accordingly, θ and $\varphi(\lambda,\sigma^2,\sigma_\varepsilon^2)$ can be jointly calibrated (Xu and Valocchi, 2015; Xu et al., 2017; Yue Pan et al., 2020). Based on Bayes' theorem, the posterior distribution can be expressed as:

$$p(\theta,\varphi|D)\propto p(D|\theta,\varphi)p(\theta)p(\varphi) \quad (3)$$

where $p(\theta)$ and $p(\varphi)$ denote the prior distributions of the physical parameters θ and the hyperparameters φ, respectively, and $p(D|\theta,\varphi)$ represents the likelihood function. In this study, a commonly used Gaussian likelihood function is adopted, and its logarithmic form can be written as:

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

$$\log(p(D|\theta,\varphi)) = -\frac{1}{2}(D-f-\mu)^T \Sigma^{-1}(D-f-\mu) - \frac{1}{2}\log|\Sigma| - \frac{N}{2}\log 2\pi \quad (4)$$

where μ denotes the mean of the systematic prediction error, Σ represents the error covariance matrix, and N is the number of observations.

Since the posterior distribution $p(\theta,\varphi|D)$ is generally intractable to evaluate analytically, Markov chain Monte Carlo (MCMC) simulation is employed to generate posterior samples of the physical parameters θ and the GPR hyperparameters φ. In this study, the Differential Evolution Adaptive Metropolis (DREAMzs) algorithm is adopted for MCMC sampling, which has been widely applied in groundwater inverse modeling and uncertainty analysis.

Based on the posterior samples of θ and φ obtained from MCMC simulation, predictions of the physical model f and the structural error model b can be generated for a new input x^*, denoted as $f^*(\theta)$ and $b^*(x^*,\varphi)$, respectively. The latter can be expressed as:

$$b^* \left| D-f, D, D^* \sim N(\overline{b^*}, C_b(b^*)) \right. \quad (5)$$

where,

$$b^* = \mu^* + C^{*T} C^{-1}(D-f-\mu) \quad (6)$$

$$C_b(b^*) = C^{**} - C^{*T} C^{-1} C^* \quad (7)$$

where $\overline{b^*}$ denotes the posterior mean, $C_b(b^*)$ represents the posterior covariance matrix, and μ^* denotes the prior mean $\mu(x^*,\varphi)$. The prior covariance matrices C^* and C^{**} are defined as $C^*_{i,j} = f(x_i, x^*_j)$ and $C^{**}_{i,j} = f(x^*_i, x^*_j)$, respectively.

The final corrected prediction D^* can be expressed as:

$$D^* = f^* + b^* + \varepsilon^* \quad (8)$$

where f^*, b^* and ε^* denote the predictions of the groundwater model, the structural error model, and the measurement error at x^*, respectively.

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

2.2. Likelihood Function-based Coupling of Physical Constraints

In practical applications, predictions of groundwater contaminant transport simulations are typically required to satisfy multiple key physical constraints, such as contaminant mass conservation and non-negativity of concentration. In this study, each physical constraint is represented by a set of functions:

$$G(y)=\left[g_1(y),g_2(y),...,g_{n_c}(y)\right]^T=[0,0,...,0]^T \tag{9}$$

$$Z(y)=\left[z_1(y),z_2(y),...,z_{n_f}(y)\right]^T=[0,0,...,0]^T \tag{10}$$

where y denotes the corrected prediction (i.e., the sum of the physical model, the structural error model, and the measurement error), which is a function of the parameters θ and φ, i.e., $y=F(\theta,\varphi)$. $g_i(y),i=1,2,...,n_c$ represent the equality constraints, where n_c denotes the number of equality constraints. $z_i(y),i=1,2,...,n_f$ represent the inequality constraints, where n_f denotes the number of inequality constraints. The zero matrix on the right-hand side indicates that the corrected predictions are required to satisfy the corresponding physical constraints at each observation time.

The equality constraints are assumed to follow a zero-mean Gaussian distribution, and Eq. (9) can therefore be expressed as:

$$G(y)=G(F(\theta,\varphi))\sim N(0,\Sigma_c) \tag{11}$$

where Σ_c is a diagonal covariance matrix used to control the strictness of constraints (Wu et al., 2019).

Therefore, a likelihood function can be constructed to quantify the extent to which the corrected predictions satisfy the equality constraints, conditional on the parameters θ and φ:

$$p(G(y)=0|\theta,\varphi)=\frac{1}{\sqrt{(2\pi)^{n_c}|\Sigma_c|}}\exp(-\frac{1}{2}G(F(\theta,\varphi))^T\Sigma_c^{-1}G(F(\theta,\varphi))) \tag{12}$$

For the inequality constraints, Eq. (10) can be transformed into the corresponding equality constraints as:

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

$$\min(0, Z(y)) = 0 \tag{13}$$

Similarly, under the assumption of a zero-mean Gaussian distribution, Eq. (13) can be written as:

$$\min(0, Z(y)) \sim N(0, \Sigma_f) \tag{14}$$

The corresponding likelihood function can be constructed as:

$$p(Z(y) \geq 0 | \theta, \varphi) = \frac{1}{\sqrt{(2\pi)^{n_f} |\Sigma_f|}} \exp(-\frac{1}{2}[\min(0, Z(y))]^T \Sigma_f^{-1} \min(0, Z(y))) \tag{15}$$

Accordingly, the posterior distribution conditional on both the observational data and multiple physical constraints (including the equality and inequality constraints described above) can be expressed as:

$$\begin{aligned} p(\theta, \varphi | D, G(y) = 0, Z(y) \geq 0) &\propto p(D, G(y) = 0, Z(y) \geq 0 | \theta, \varphi) p(\theta) p(\varphi) \\ &= p(D | \theta, \varphi) p(G(y) = 0 | \theta, \varphi) p(Z(y) \geq 0 | \theta, \varphi) p(\theta) p(\varphi) \end{aligned} \tag{16}$$

where $p(D|\theta, \varphi)$ corresponds to the likelihood function that quantifies the fit between the predictions and the observational data, as given in Eq. (4), $p(G(y) = 0|\theta, \varphi)$ and $p(Z(y) \geq 0|\theta, \varphi)$ correspond to the likelihood functions incorporating the equality and inequality constraints defined in Eqs. (12) and (15), respectively.

This formulation can be further generalized, and Eq. (16) is equally applicable when additional physical constraints are incorporated. In this case, the total log-likelihood function can be expressed as the sum of the log-likelihoods of the individual sub-likelihood functions:

$$L_T = L_1 + L_2 + L_3 ... + L_n \tag{17}$$

where L_T denotes the total likelihood, L_1 represents the likelihood associated with the predictions, and $L_2, L_3, \ldots, L_n$ denote the likelihood functions corresponding to different physical constraints.

Therefore, MCMC sampling is performed based on the combined likelihood function (Eq. 17) to jointly identify the posterior distributions of the groundwater model parameters θ and the GPR hyperparameters φ.

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

2.3. Dynamic Weight-updating Strategy with a Stopping Criterion

Because different sub-likelihood functions correspond to distinct physical meanings, their magnitudes may differ substantially, leading to unequal competition among the sub-likelihood functions. Therefore, it is necessary to assign weights to the sub-likelihood functions and dynamically update them according to their relative magnitudes. In this study, an inverse-magnitude weighting method is adopted to weight the sub-likelihood functions as follows (Guger et al., 2020; Li et al., 2018):

$$L_{T,i} = w_{1,i} \bullet L_{1,i} + w_{2,i} \bullet L_{2,i} + \ldots + w_{n,i} \bullet L_{n,i} \tag{18}$$

where $L_{1,i}, L_{2,i}, \ldots, L_{n,i}$ denote the n sub-likelihood functions obtained at the i-th iteration of the Bayesian updating process under the scenario with multiple physical constraints, and $w_{1,i}, w_{2,i}, \ldots, w_{n,i}$ denote the corresponding weights. The dynamic weight-updating procedure is described as follows:

First, the mean of each sub-likelihood function is calculated as follows:

$$L_{avg,i} = \frac{\sum_{j=1}^{n} L_{j,i}}{n} \tag{19}$$

Then, the initial unnormalized weights are computed:

$$w'_{j,i} = \frac{L_{avg,i}}{L_{j,i}}, j = 1, 2, \ldots, n \tag{20}$$

Finally, the weights are normalized:

$$w_{j,i} = \frac{w'_{j,i}}{\sum_{k=1}^{n} w'_{k,i}}, j = 1, 2, \ldots, n \tag{21}$$

During the dynamic weight-updating process, determining when to terminate weight updating is crucial for ensuring the stability and convergence of MCMC sampling. In this study, a stopping criterion derived from multi-objective evolutionary algorithms is introduced (Abu Doush et al., 2023; Kadhar and Baskar, 2018), whereby weight updating is terminated when the variance of weight changes remains stable within a specified iteration window. Because the MCMC simulation involves multiple chains, the stopping criterion must account for weight updates across different Markov

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

chains. The procedure is as follows:

(1) After the *i*-th iteration, compute the weight $w_{j,i}^{r}$ of the *j*-th sub-likelihood function for the *r*-th chain.

(2) Compute the change in each weight between two consecutive iterations $\Delta w_{j,i}^{r} = \left| w_{j,i}^{r} - w_{j,i-1}^{r} \right|$.

(3) Over the most recent γ iterations, compute the variance of the weight changes for each sub-likelihood function and each chain as $V_{j,\gamma}^{r} = Variance(\Delta w_{j,i-\gamma+1}^{r}, \Delta w_{j,i-\gamma+2}^{r}, ..., \Delta w_{j,i}^{r})$. The maximum value $MV_{j^*,\gamma}^{r^*}$ is then recorded, and the corresponding weight change is denoted as $M\Delta w_{j^*,i}^{r^*}$, which is then used in the subsequent χ^2 test.

(4) A χ^2 test is conducted to statistically assess whether $M\Delta w_{j^*,i}^{r^*}$ over the recent γ iterations is smaller than the prescribed tolerance variance (*Tol_fun*). The test statistic is given by:

$$chi(M\Delta w_{j^*,i}^{r^*}) = \frac{Variance(\Delta w_{j^*,i-\gamma+1}^{r^*}, \Delta w_{j^*,i-\gamma+2}^{r^*}, ..., \Delta w_{j^*,i}^{r^*})(\gamma - 1)}{Tol_{fun}^{2}} \tag{22}$$

The corresponding probability is:

$$P(M\Delta w_{j^*,i}^{r^*}) = \chi^2 test[chi(M\Delta w_{j^*,i}^{r^*}), (\gamma - 1)] \tag{23}$$

When $P(M\Delta w_{j^*,i}^{r^*}) > 99\%$, the weight changes are considered statistically insignificant, and the weight-updating process is terminated.

In summary, the overall workflow of the proposed framework for structural error correction coupled with multiple physical constraints is illustrated in Fig. 1.

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

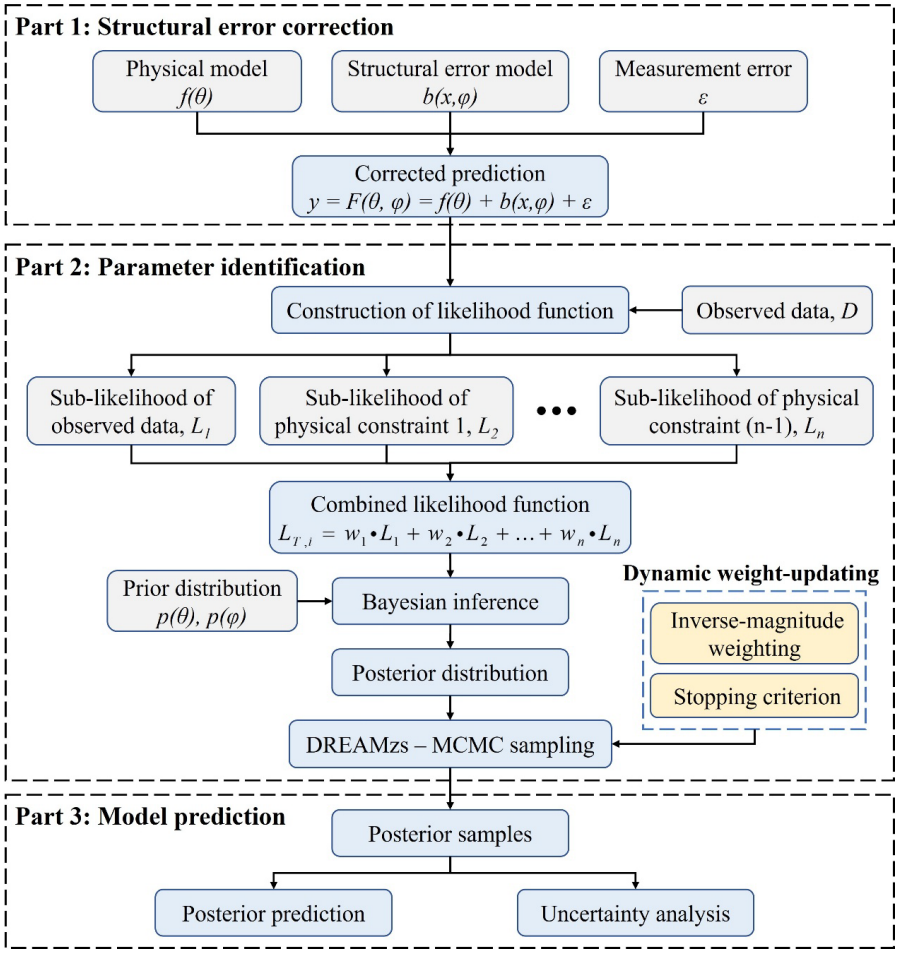

Figure 1. Overall workflow of the proposed framework.

3. Descriptions of the Two Case Studies

Two case studies were used in this study: A synthetic three-dimensional PCE reactive transport simulation and a cadmium–phosphate cotransport sand column experiment. In each case, the constructed groundwater model represents a structural simplification of the real groundwater system and therefore exhibits significant structural bias.

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

3.1. Synthetic Groundwater PCE Reactive Transport Model

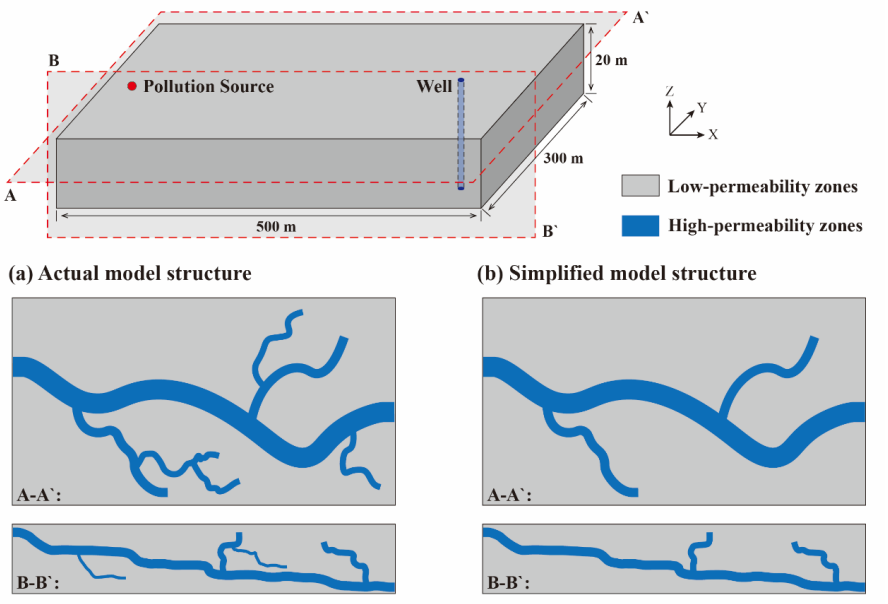

Figure 2. Schematic illustration for the synthetic groundwater PCE reactive transport model.

Case study 1 constructed a synthetic three-dimensional groundwater model to simulate the reactive transport of PCE. As illustrated in Fig. 2, the aquifer is 500 m in length, 300 m in width, and 20 m in thickness. The left and right boundaries are specified-head boundaries corresponding to a hydraulic gradient of 0.12, whereas the remaining boundaries are assigned as no-flow boundaries. The aquifer medium is horizontally isotropic, with a vertical hydraulic conductivity anisotropy of 3 and a porosity of 0.3. The overall permeability of the aquifer is relatively low, while dendritic high-permeability preferential flow pathways are distributed within it (Fig. 2a).

A PCE leakage source is located near the left boundary at the top of the aquifer (x = 75 m, y = 150 m, z = 20 m), with a release rate of 25 kg/d. During transport, PCE undergoes sequential reductive dechlorination to form trichloroethylene (TCE), dichloroethylene (DCE), and vinyl chloride (VC). A pumping well is located near the right boundary at the bottom of the aquifer (x = 475 m, y = 150 m, z = 7 m) with a pumping rate of 30 m^3/d, where the concentrations of PCE, TCE, DCE, and VC are

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

continuously monitored.

In practice, detailed structural information of the aquifer is often unavailable, and simplification of the aquifer structure is therefore typically required. In particular, for preferential pathways with hierarchical branching structures, smaller branch channels are often neglected. In this case, only the main preferential pathways are retained (Fig. 2b), and the PCE reactive transport model is constructed based on the simplified aquifer structure. The model is numerically solved using MODFLOW-2005 and RT3D. Gaussian process regression (GPR) is further employed to correct the structural error introduced by simplification of high-permeability preferential pathways.

Except for the representation of the preferential pathways, the simplified model is identical to the actual model in all other respects. The total simulation period is 500 days with a time step of 10 days. The parameter settings of the actual model are listed in Table 1. The contaminant concentrations simulated at the pumping well are treated as observational data. Gaussian white noise with a mean of 0 and a variance of 0.01 is added to the simulated concentrations, yielding a total of 50 concentration observations, of which the first 30 are used for parameter identification and the remaining 20 for model validation.

Table 1. Parameter configuration for the synthetic groundwater PCE reactive transport model with the actual model structure.

Parameters	Unit	Value
Porosity (ϕ)	-	0.3
Horizontal anisotropy of hydraulic conductivity (A_{Kh})	-	1.0
Vertical anisotropy of hydraulic conductivity (A_{Kv})	-	3.0
Horizontal anisotropy of dispersivity (A_{Dh})	-	0.3
Vertical anisotropy of dispersivity (A_{Dv})	-	0.3
Horizontal hydraulic conductivity of low-permeability zones (K_l)	$m \cdot d^{-1}$	1.5
Horizontal hydraulic conductivity of high-permeability zones (K_h)	$m \cdot d^{-1}$	8.0
Longitudinal dispersivity of low-permeability zones (D_l)	m	1.5
Longitudinal dispersivity of high-permeability zones (D_h)	m	8

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

3.2. Sand Column Cadmium-phosphate Cotransport Model

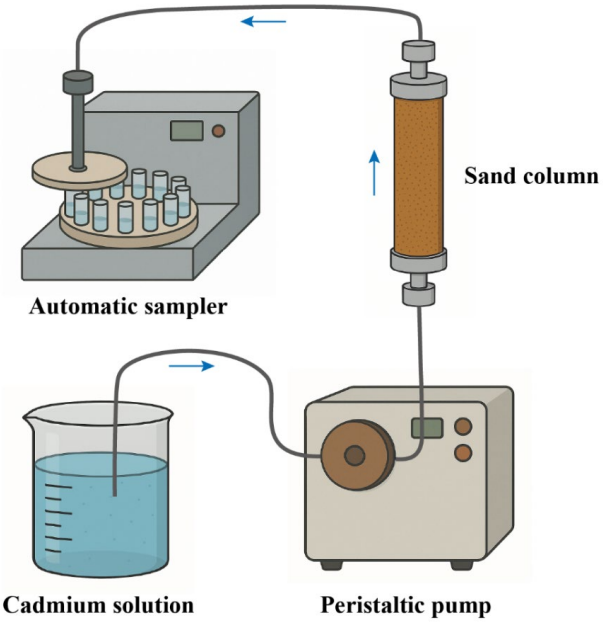

Figure 3. Schematic illustration of the sand column Cd-phosphate cotransport experiment.

To investigate the influence of phosphate on cadmium (Cd) transport in natural groundwater environments, a one-dimensional cadmium–phosphate cotransport sand column experiment was conducted (Wu et al., 2022) (Fig. 3). A polypropylene column with a height of 5 cm and an inner diameter of 1.5 cm was used. Stainless steel screens with a pore size of 10 μm were placed at both ends of the column to support the porous medium. The column was uniformly packed with dry soil, and the average porosity (0.36) and bulk density (1.69 g/cm^3) were determined using the gravimetric method. A peristaltic pump was used to drive the solution upward through the column at a flow rate of 0.2 mL/min. After complete saturation, the column was flushed with background solution for 12 h, followed by a tracer test to determine flow and solute transport parameters. Once geochemical conditions stabilized, 420 mL of a mixed solution containing 0.1 mM Cd and 1 mM phosphate was injected into the column using the

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

peristaltic pump. After completion of the mixed-solution injection, the influent was switched to background solution without Cd. During the experiment, effluent samples were collected every 90 min using an automatic fraction collector, and Cd concentrations were measured, yielding a total of 58 concentration observations.

In real-world groundwater contamination scenarios, the spatial distribution of phosphate is often uncertain and its influence on Cd transport is therefore commonly neglected. Accordingly, case study 2 does not account for the cotransport mechanism between Cd and phosphate. A one-dimensional groundwater model for Cd transport was developed using HYDRUS-1D (Nkedikizza et al., 1984; Vengenuchten, 1980), in which a two-site sorption model coupled with a Freundlich isotherm was employed to describe Cd transport. The parameter settings of the physical model are listed in Table 2. In addition, Gaussian process regression (GPR) was employed to correct the structural error arising from neglecting the Cd-phosphate cotransport mechanism. Of the 58 observations, the first 32 were used for parameter identification and the remaining 26 for validation.

Table 2. Parameter configuration of the sand column Cd-phosphate cotransport model.

Parameters	Unit	Value
Residual Water Content (θ_r)	$cm^3 \cdot cm^{-3}$	0.0
Saturated Water Content (θ_s)	$cm^3 \cdot cm^{-3}$	0.4
Saturated Hydraulic Conductivity (K_s)	$cm \cdot min^{-1}$	0.1132
Pore connectivity factor (P_L)	-	0.5
Bulk density (R_O)	$g \cdot cm^{-3}$	1.24
Longitudinal dispersivity (D_v)	cm	0.74

4. Application of the Methods

In each case study, contaminant transport simulations are first performed using a groundwater model with structural error, after which the model is corrected within the DDM framework. In addition, three scenarios are considered: without physical constraints, with a single physical constraint, and with multiple physical constraints.

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

4.1. DDM for Synthetic Groundwater PCE Reactive Transport Model

As described in Section 3.1, case study 1 constructed a groundwater model with a simplified representation of preferential pathways in the aquifer to simulate PCE reactive transport. In this model, the hydraulic conductivities and dispersivities of the low- and high-permeability zones are treated as random variables with uniform prior distributions. Meanwhile, the GPR hyperparameters σ, λ and σ_ε are assigned exponential, gamma, and uniform priors, respectively. Table 3 summarizes the prior information for the physical parameters and the hyperparameters.

Table 3. Prior distributions of parameters for the synthetic groundwater PCE reactive transport model.

Parameter	Ranges	Type
K_l [m/d]	[1.0, 5.0]	Uniform
K_h [m/d]	[5.0, 12.0]	Uniform
D_l [m]	[1.0, 5.0]	Uniform
D_h [m]	[5.0, 12.0]	Uniform
σ	[0.001, 0.16]	Exponential
λ	[0.1, 3.0]	Gamma
σ_ε	[0.001, 0.12]	Uniform

In this case study, because carbon and chlorine are the key constituent elements of PCE and its dechlorination products, carbon and chlorine mass conservation constraints are incorporated into the DDM. Under the single-constraint scenario, only the carbon mass conservation constraint is coupled, whereas under the multiple-constraint scenario both carbon and chlorine mass conservation constraints are coupled simultaneously. For the entire system, the total mass of leaked carbon/chlorine is equal to the total carbon/chlorine mass contained in PCE, TCE, DCE, and VC. Moreover, at any time, the total mass of leaked carbon/chlorine m_l should equal the sum of three components: the mass remaining in the aquifer m_r, the mass leaving the aquifer through its boundaries m_b, and the mass extracted by the pumping well m_w. The mass balance error is defined as M_{error}, and the corresponding mass conservation constraints for carbon/chlorine can be expressed as:

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

$$M_{error} = m_l - (m_r + m_b + m_w) = 0 \tag{24}$$

Since the observation intervals are much larger than the simulation time step of the physical model, the physical constraints at different observation times are assumed to be independent and enforced with equal strength. The confidence on the accuracy of each physical constraint ($\sigma_{c,i}^2$) is set to 0.2. Accordingly, the likelihood function associated with the carbon/chlorine mass conservation constraints can be expressed as:

$$\begin{aligned} & p(M_{error} = 0 \mid \theta, \varphi) \\ & = \frac{1}{\sqrt{(2\pi)^n \left|\sum_c\right|}} \exp\left(-\frac{1}{2}[m_l - (m_r + m_b + m_w)]^T \sum_c^{-1} [m_l - (m_r + m_b + m_w)]\right) \end{aligned} \tag{25}$$

In addition, following Kadhar and Baskar (2018), the sliding window length used for dynamic weight updating is set to $\gamma = 10$, and the minimum tolerance variance is set to $Tol\_fun = 10^{-4}$. The dynamic sub-likelihood weight-updating strategy with a stopping criterion is implemented as described in Section 2.3. Three scenarios are considered: without physical constraints, with only the carbon mass conservation constraint, and with both carbon and chlorine mass conservation constraints.

After specifying the prior distributions and constructing the combined likelihood function, MCMC simulations were conducted to infer the posterior distributions of the parameters. Both unconstrained and constrained DDM configurations use the same MCMC settings based on the DREAMzs algorithm, with three parallel Markov chains, each running for 9000 iterations, and the first 3000 iterations designated as burn-in. Finally, based on the identified posterior parameter distributions, the predictive performance of the DDM for the concentrations of the four chlorinated hydrocarbons was evaluated under the three scenarios.

4.2. DDM for Sand Column Cd-phosphate Cotransport Model

As described in Section 3.2, case study 2 developed a simplified Cd transport model that does not consider the cadmium-phosphate cotransport mechanism. The proportion coefficient of equilibrium sorption *f*, the Freundlich isotherm coefficient K_F,

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

and the isotherm shape parameter β of the physical model are treated as random variables with uniform prior distributions. The GPR hyperparameters σ, λ and σ_ε are assigned exponential, gamma, and uniform prior distributions, respectively. The prior information for the physical model parameters and the GPR hyperparameters is summarized in Table 4.

Table 4. Prior distributions of parameters for the sand column Cd-phosphate cotransport model.

Parameter	Ranges	Type
f	[0.4, 0.6]	Uniform
K_F [cm^3/mg]	[30.0, 40.0]	Uniform
β	[0.6, 1.2]	Uniform
σ	[0.01, 0.11]	Exponential
λ	[0.1, 1.0]	Gamma
σ_ε	[0.0001, 0.01]	Uniform

In this case study, Cd mass conservation and Cd concentration bound constraints are incorporated into the DDM. Under the single-constraint scenario, only the Cd mass conservation constraint is coupled. At any given time t_i, the total mass of Cd entering the column is assumed to be equal to the sum of the Cd masses present in different states. Specifically, before Cd breakthrough ($t_i \le t_0$), the total mass consists of the dissolved-phase mass m_d and the mass adsorbed onto the solid phase m_a. After breakthrough ($t_i > t_0$), the total mass additionally includes the Cd mass in the effluent, m_{out}. Accordingly, the Cd mass conservation constraint is defined as:

$$M_{error,t_i} = \begin{cases} m_{in,t_i} - (m_{d,t_i} + m_{a,t_i}) = 0, t_i \le t_0 \\ m_{in,t_i} - (m_{d,t_i} + m_{a,t_i} + m_{out,t_i}) = 0, t_i \le t_0 \end{cases} \quad (26)$$

where m_{in,t_i} denotes the total Cd mass entering the column; m_{d,t_i}, m_{a,t_i} and m_{out,t_i} denote the Cd mass in the dissolved phase, sorbed phase, and effluent, respectively; and t_0 represents the time of first Cd breakthrough.

Consistent with case study 1, the confidence on the accuracy of each physical constraint ($\sigma_{c,i}^2$) is set to 0.2, and the corresponding likelihood function can be written as:

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

$$p(M_{error}=0|\theta,\varphi) = \begin{cases} \frac{1}{\sqrt{(2\pi)^{n_0}|\Sigma_c|}}\exp(-\frac{1}{2}[m_{in}-(m_d+m_a)]^T\Sigma_c^{-1}[m_{in}-(m_d+m_a)]) \\ \frac{1}{\sqrt{(2\pi)^{(n-n_0)}|\Sigma_c|}}\exp(-\frac{1}{2}[m_{in}-(m_d+m_a+m_{out})]^T\Sigma_c^{-1}[m_{in}-(m_d+m_a+m_{out})]) \end{cases} \quad (27)$$

where n is the total number of observations and n_0 is the number of observations before Cd breakthrough.

Under the multiple-constraint scenario, the Cd concentration bound constraint is further coupled. The experimental breakthrough curve indicates that the upper bound of Cd concentration does not exceed 0.75 mg/cm³. Therefore, the Cd concentration bound constraint is defined as:

$$0 \le C_{Cd} \le 0.75 \quad (28)$$

This constraint is transformed into an equivalent equality constraint and assumed to follow a Gaussian distribution $N(0,\Sigma_f)$, which can be expressed as:

$$Z(y) = [\max(0,-C_{Cd}) + \max(0,C_{Cd}-0.75)] = 0 \quad (29)$$

Accordingly, the likelihood function associated with the Cd concentration bound constraint can be expressed as:

$$p(Z(y)=0|\theta,\varphi) = \frac{1}{\sqrt{(2\pi)^n|\Sigma_f|}}\exp(-\frac{1}{2}[Z(y)=0]^T\Sigma_c^{-1}[Z(y)=0]) \quad (30)$$

As in case study 1, the sliding window length ($\gamma=10$) and minimum tolerance variance ($Tol\_fun=10^{-4}$) are used for dynamic sub-likelihood weight updating and the stopping criterion. As in case study 1, three scenarios are considered: without physical constraints, with only the Cd mass conservation constraint, and with both the Cd mass conservation and Cd concentration bound constraints.

MCMC simulations were then conducted to infer the posterior distributions of the parameters. Under all three scenarios, identical MCMC settings based on the DREAMzs algorithm are employed, with three parallel Markov chains, each running for 9000 iterations and the first 3000 iterations designated as burn-in. Based on the

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

identified posterior parameter distributions, the predictive performance of the DDM for Cd concentrations is evaluated under the three scenarios.

5. Results and Discussions

5.1. Parameter Estimations

5.1.1. Synthetic Groundwater PCE Reactive Transport Model

Fig. 4 presents the posterior distributions of the PCE reactive transport model parameters under the three scenarios, where the first row shows the physical model parameters and the second row showing the GPR hyperparameters. The results indicate that both the presence and the number of coupled physical constraints significantly influence the posterior distributions. Under the unconstrained and single-constraint scenarios, the posterior distributions are generally similar, exhibiting relatively wide ranges and lower peak probability densities, indicating greater parameter uncertainty. In contrast, when multiple physical constraints are coupled simultaneously, the posterior distributions contract substantially. Specifically, the posterior ranges become substantially narrower and the peak densities increase markedly. Notably, the posterior distributions of the physical model parameter K_l and the GPR hyperparameters σ, λ and σ_ε show particularly pronounced changes in terms of range, mean value, and peak density (Figs. 4a and 4e-4g). In addition, the physical model parameters K_h, L_l and L_h exhibit more pronounced multimodal characteristics (Figs. 4b-4d), suggesting the presence of multiple plausible parameter combinations.

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

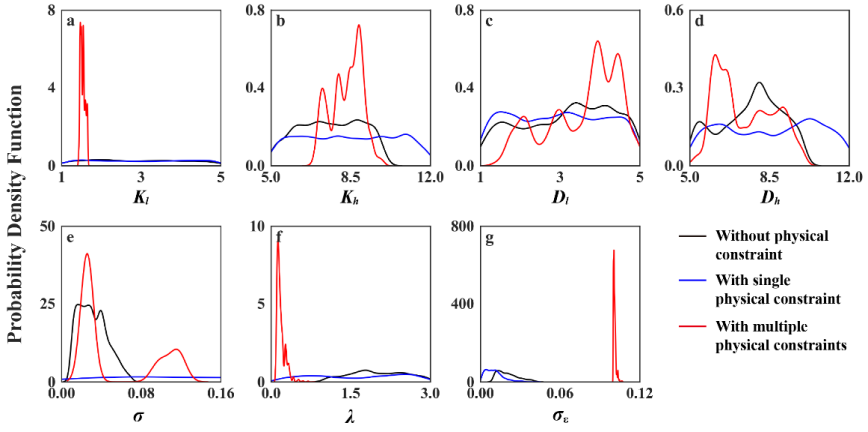

Figure 4. Parameter posterior distributions for the synthetic groundwater PCE reactive transport model under the three scenarios.

5.1.2. Sand Column Cd-phosphate Cotransport Model

Fig. 5 presents the posterior distributions of the parameters of the Cd-phosphate cotransport model under the three scenarios, where the first and second rows correspond to the physical model parameters and the GPR hyperparameters, respectively. Under the unconstrained and single-constraint scenarios, most parameters exhibit wide posterior ranges and low peak probability densities, indicating greater parameter uncertainty. Except for the relatively flat probability density curves for the physical model parameter K_F and the GPR hyperparameter λ (Figs. 5b and 5e), the posterior distributions of the remaining parameters display varying degrees of multimodality. In contrast, when multiple physical constraints are simultaneously coupled, the posterior distributions of all parameters exhibit substantial contraction and markedly higher peak densities. Among them, the changes are particularly pronounced for the physical model parameter K_F and the GPR hyperparameters σ and λ (Figs. 5b and 5d-5e).

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

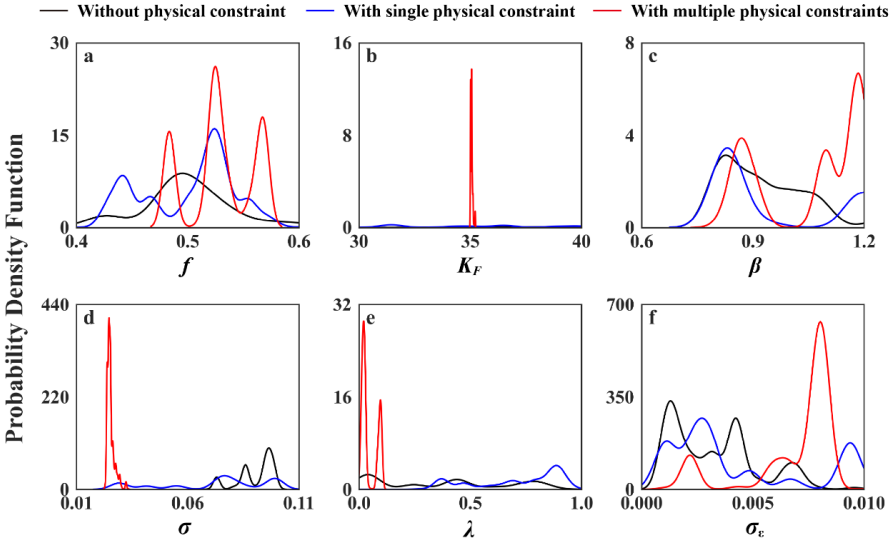

Figure 5. Parameter posterior distributions for the sand column Cd–phosphate cotransport model under the three scenarios.

5.2. Mass Conservation Evaluation

5.2.1. Synthetic Groundwater PCE Reactive Transport Model

Fig. 6 illustrates the temporal variation in the carbon and chlorine mass balance error rates (ERs) predicted by the DDM under the three scenarios for case study 1. During the simulation, the contamination source continuously releases PCE at a rate of 25 kg/d, containing 3.63 kg of carbon and 21.37 kg of chlorine. The mass balance error rate (ER) for carbon/chlorine is defined as:

$$ER = \frac{M_T - M_S}{M_T} \times 100\% \tag{31}$$

where M_T and M_S denote the true total mass of carbon/chlorine in the system and the corresponding total mass predicted by the DDM, respectively.

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

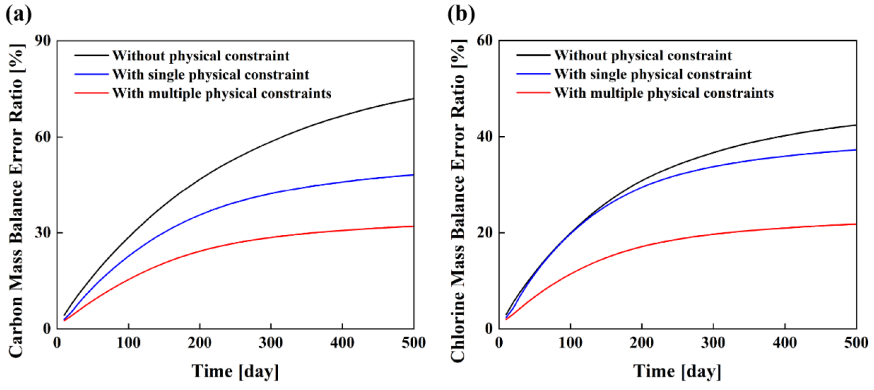

Figure 6. Variation in the carbon/chlorine mass balance error rate for the synthetic groundwater PCE reactive transport model under the three scenarios.

The results indicate that coupling physical constraints significantly reduces the mass conservation error of the DDM predictions. Under the unconstrained scenario, the mean mass conservation ERs for carbon and chlorine reach 48.62% and 30.65%, corresponding to average mass errors of 1.76 kg and 6.55 kg, respectively. After coupling the carbon mass conservation constraint, the mean ER for carbon decreases to 35.03%, with the average mass error reduced to 1.27 kg. Meanwhile, the mean ER for chlorine decreases slightly to 28.27%, with an average mass error of 6.04 kg. When both carbon and chlorine mass conservation constraints are simultaneously coupled, the mean ERs for carbon and chlorine further decrease to 23.63% and 16.49%, with the corresponding average mass errors reduced to 0.86 kg and 3.52 kg, respectively. Moreover, compared with the unconstrained scenario, coupling physical constraints results in a more gradual increase in the carbon/chlorine ER over time, followed by stabilization in the later stages of the simulation.

5.2.2. Sand Column Cd-phosphate Cotransport Model

Fig. 7 presents the temporal variation of the Cd mass conservation error rate (ER) predicted by the DDM under the three scenarios in case study 2. Over the entire experiment period, a total of 4.72 mg of Cd is introduced into the soil column. Under the unconstrained scenario, the mean mass conservation ER for Cd is 30.14%, with an average mass error of 1.42 mg. The peak ER reaches 57.14%, corresponding to a

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

maximum mass error of 2.70 mg. When only the Cd mass conservation constraint is coupled, the mean and peak ERs decrease markedly to 13.64% and 26.07%, respectively, with the corresponding average and maximum mass errors substantially reduced to 0.64 mg and 1.23 mg. When both the Cd mass conservation and concentration bound constraints are simultaneously coupled, the mean and peak ERs further decrease to 10.36% and 19.33%, respectively, with the corresponding average and maximum mass errors reduced to 0.49 mg and 0.91 mg.

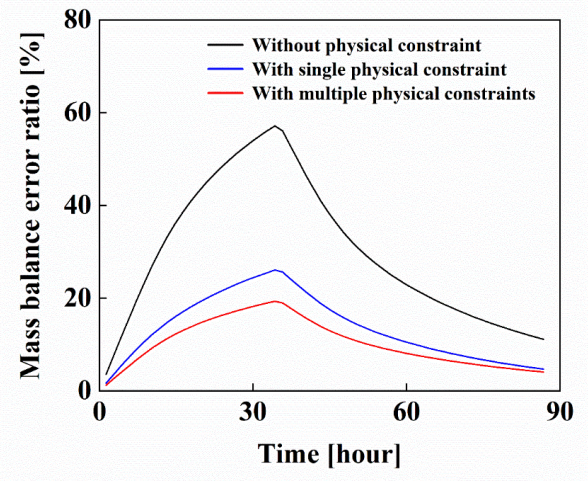

Figure 7. Variation in the cadmium mass balance error rate for the sand column Cd-phosphate cotransport model under the three scenarios.

5.3. Prediction Performance

5.3.1. Synthetic Groundwater PCE Reactive Transport Model

Fig. 8 presents the temporal variations in concentrations of four chlorinated hydrocarbons predicted by the DDM under the three scenarios in case study 1. Under the unconstrained scenario, the model captures the overall trends of chlorinated hydrocarbon concentrations. However, noticeable discrepancies remain between the posterior means and the observations during the validation period (Figs. 8a-8d), particularly for DCE and VC (Figs. 8c and 8d). After coupling a single physical constraint, the DDM shows improved predictive performance. The predicted TCE concentrations become closer to the observations, whereas substantial deviations from

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

the observations remain for the other chlorinated hydrocarbons during the validation period. In contrast, when multiple physical constraints are simultaneously coupled, the predictive accuracy for all four chlorinated hydrocarbons improves markedly, with particularly pronounced improvements during the validation period (Figs. 8i-8l).

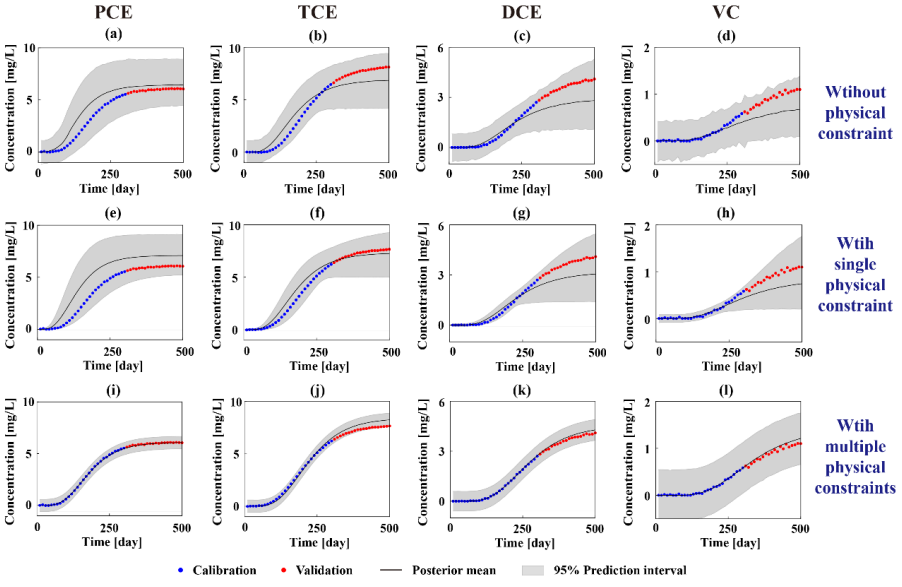

Figure 8. DDM predictions of chlorinated ethenes concentrations under the three scenarios.

Moreover, under the unconstrained scenario, the 95% prediction intervals are generally wider, especially during the validation period (Figs. 8a-8d). When a single physical constraint is coupled, the predictive intervals narrow during the calibration period but remain wide during the validation period, particularly for DCE and VC (Figs. 8g and 8h). In contrast, simultaneously coupling multiple physical constraints markedly reduces the prediction intervals for all chlorinated hydrocarbons (Figs. 8i-8l), indicating a substantial reduction in predictive uncertainty.

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

Table 5. Evaluation of predictions for the synthetic groundwater PCE reactive transport model under the three scenarios.

Constraints	Calibration/ Validation	Species	NSE	MAE	RMSE
Without physical constraint	Calibration	PCE	0.7300	0.9444	1.0728
		TCE	0.9129	0.15615	0.6978
		DCE	0.9410	0.1709	0.2313
		VC	0.8732	0.0427	0.0708
	Validation	PCE	-13.9883	0.4076	0.4111
		TCE	-0.6106	1.10108	1.0313
		DCE	-8.5719	1.0405	1.0561
		VC	-4.2188	0.3369	0.3467
With single physical constraint	Calibration	PCE	0.4866	1.3393	1.4795
		TCE	0.9282	0.5057	0.6043
		DCE	0.9642	0.1383	0.1801
		VC	0.9230	0.0351	0.0552
	Validation	PCE	-96.8477	1.0494	1.0505
		TCE	0.2689	0.2481	0.2762
		DCE	-4.9220	0.8148	0.8307
		VC	-2.5998	0.2777	0.2879
With multiple physical constraints	Calibration	PCE	0.9994	0.0417	0.0496
		TCE	0.9955	0.1241	0.1505
		DCE	0.9991	0.0216	0.0285
		VC	0.9964	0.0091	0.0119
	Validation	PCE	0.8126	0.0356	0.0460
		TCE	0.6444	0.4863	0.4946
		DCE	0.7920	0.1471	0.1557
		VC	0.7728	0.0655	0.0723

Table 5 summarizes the quantitative evaluation of DDM predictive performance under the three scenarios in case study 1. Compared with the unconstrained scenario, coupling a single physical constraint leads to a modest improvement in the NSE of the predicted concentrations for the four chlorinated hydrocarbons during the validation period, although the improvement remains limited. Except for PCE, for which MAE and RMSE increase slightly, MAE and RMSE decrease for the other chlorinated hydrocarbons. Specifically, the MAE and RMSE of TCE decrease by 77.47% and 73.22%, respectively, whereas those of DCE and VC decrease by only about 20%. When multiple physical constraints are coupled, the predictive performance of the

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

DDM improves markedly relative to the single-constraint scenario. During the validation period, TCE behaves differently from the other chlorinated hydrocarbons: its NSE increases from 0.2689 to 0.6444 (approximately 139.64%), whereas its MAE and RMSE increase slightly. In contrast, the NSE values of the other chlorinated hydrocarbons increase markedly from negative values to approximately 0.8, accompanied by substantial reductions in MAE and RMSE. Specifically, both MAE and RMSE decrease by approximately 95% for PCE, 81% for DCE, and 75% for VC.

5.3.2. Sand Column Cd-phosphate Cotransport Model

Fig. 9 presents the temporal variation in Cd concentration predicted by the DDM under the three scenarios in case study 2. Under the unconstrained scenario, the model captures the overall trend of Cd concentration, which first increases and then decreases. However, noticeable discrepancies remain between the posterior means and the observations, with the concentration peak overestimated and concentrations during the validation period markedly underestimated (Fig. 9a). After coupling the Cd mass conservation constraint, the DDM predictions become more consistent with the observations. Except for residual deviations at the early and late stages of the simulation, the posterior mean aligns more closely with the observations at most time points (Fig. 9b). When multiple physical constraints are simultaneously coupled, the predictive accuracy of the DDM improves markedly, with particularly evident improvements at the concentration peak and during the validation period (Fig. 9c).

In addition, predictive uncertainty differs substantially among the scenarios. Under the unconstrained scenario, the 95% prediction intervals are relatively wide (Fig. 9a). After coupling a single physical constraint, the width of the prediction intervals changes only marginally. Notably, some posterior predictions still fall outside the physically plausible range of Cd concentration, with the predicted maximum exceeding the highest observed value and the predicted minimum concentration dropping below zero (Fig. 9b). In contrast, when the Cd concentration bound constraint is additionally coupled, all posterior predictions remain within the physically admissible range, and the prediction intervals become substantially narrower, indicating a marked reduction in

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

predictive uncertainty.

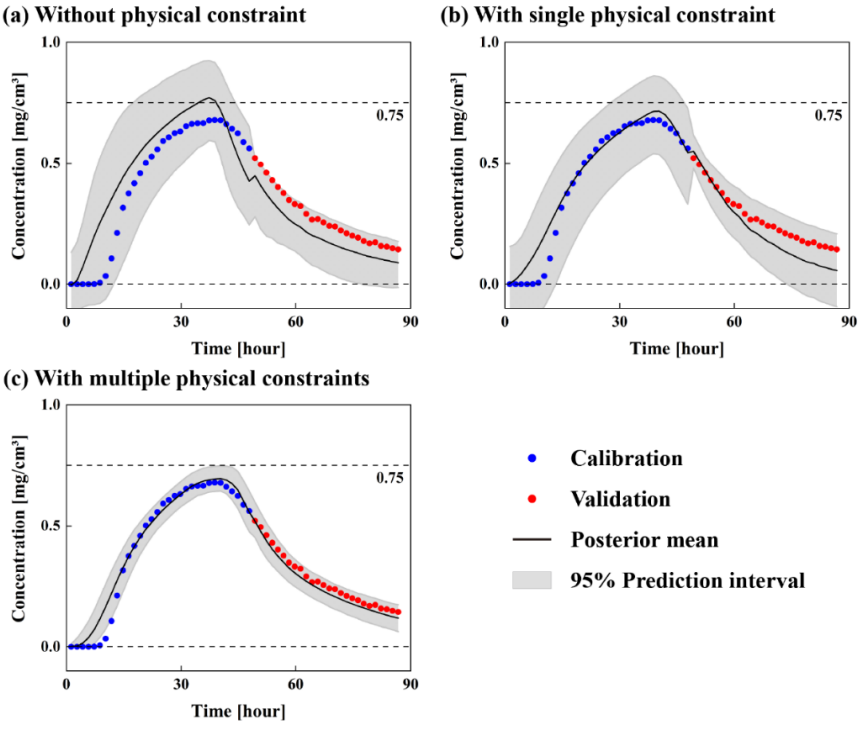

Figure 9. DDM predictions of cadmium concentration under the three scenarios.

Table 6 presents the quantitative evaluation of the DDM predictive performance under the three scenarios in case study 2. Compared with the unconstrained scenario, coupling a single physical constraint increases the NSE of Cd concentration predictions during the validation period from 0.4874 to 0.6241 (an improvement of approximately 28.05%), while the MAE and RMSE decrease by approximately 22.91% and 14.33%, respectively. When multiple physical constraints are simultaneously coupled, the NSE during the validation period further increases from 0.6241 to 0.9419 (an improvement of approximately 50.92%) relative to the single-constraint scenario, while the MAE and RMSE decrease by approximately 56.52% and 60.61%, respectively, indicating a substantial improvement in predictive performance.

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

Table 6. Evaluation of predictions for the sand column Cd-phosphate cotransport model under the three scenarios.

Constraints	Calibration/ Validation	NSE	MAE	RMSE
Without physical constraint	Calibration	0.7696	0.1018	0.1198
	Validation	0.4874	0.0716	0.0726
With single physical constraint	Calibration	0.9460	0.0381	0.0580
	Validation	0.6241	0.0552	0.0622
With multiple physical constraints	Calibration	0.9704	0.0269	0.0429
	Validation	0.9419	0.0240	0.0245

5.4. Discussions

5.4.1. Updating of Sub-likelihood Function Weights

Since sub-likelihood functions associated with different physical constraints represent distinct physical mechanisms, their numerical scales often differ by several orders of magnitude. In case study 1, sub-likelihood values differ by up to six orders of magnitude, whereas in case study 2 they differ by approximately one order of magnitude. Such scale disparities cause smaller sub-likelihoods to be overshadowed by larger ones during parameter identification, thereby diminishing the influence of certain physical constraints in Bayesian inference and making it difficult for the corrected predictions to simultaneously satisfy all constraints. Therefore, dynamically updating sub-likelihood weights in the combined likelihood function is essential to ensure appropriate contributions from each sub-likelihood during parameter identification.

In both case studies, the evolution of sub-likelihood weights over the MCMC iterations indicates that the proposed dynamic weight-updating strategy with a stopping criterion plays a critical role in Bayesian inference (Figs. 10 and 11). After substantial adaptive adjustment in the early iterations, the weights rapidly converge to stable values, effectively rescaling sub-likelihoods that initially differed by several orders of magnitude to a comparable scale. This ensures that all sub-likelihoods exert meaningful influence on parameter identification while avoiding the subjectivity inherent in manually assigned weights. Notably, the stable weighting patterns obtained in the two case studies differ markedly, suggesting that fixed weights are unlikely to be suitable

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

across varying scenarios. In contrast, the proposed dynamic weight-updating strategy is well suited to combined likelihood functions that integrate multiple sub-likelihoods associated with distinct physical mechanisms. It maintains a balance between multiple physical constraints and observational information, thereby enhancing the robustness of parameter identification and the effective enforcement of physical constraints.

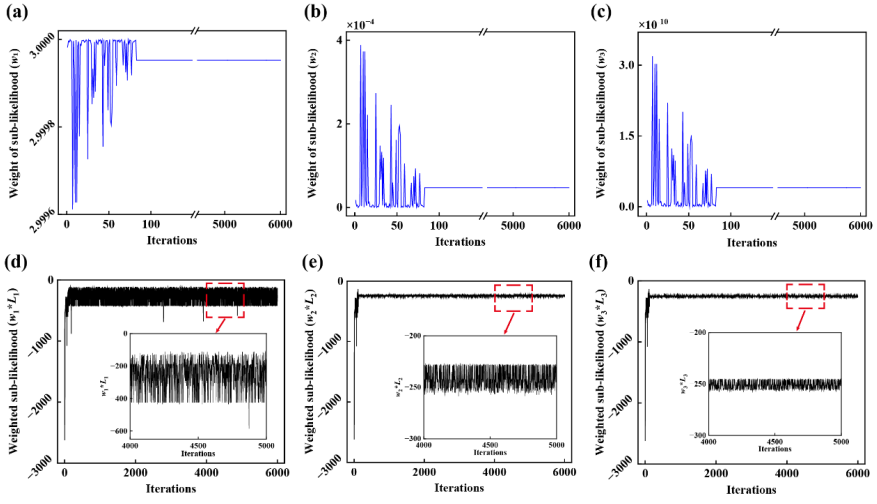

Figure 10. Dynamic updating of sub-likelihood weights (a-c) and weighted sub-likelihoods (d-f) for the synthetic groundwater PCE reactive transport model.

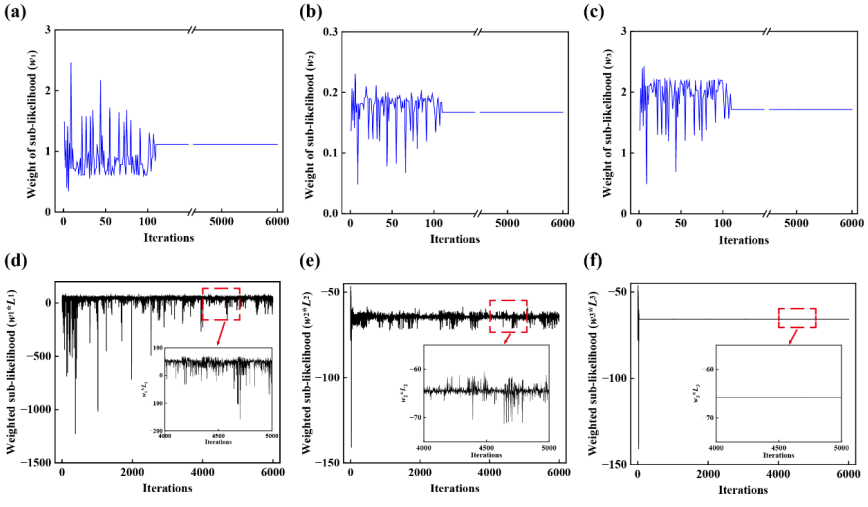

Figure 11. Dynamic updating of sub-likelihood weights (a-c) and weighted sub-likelihoods (d-f) for the sand column Cd-phosphate cotransport model.

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

5.4.2. Impact of Physical Constraints on Posterior Model Parameters and Predictions

The presence and the number of coupled physical constraints significantly influence the posterior distributions of DDM parameters (Figs. 4 and 5). Without physical constraints, the structural error correction process lacks the constraints imposed by physical mechanisms, causing both the physical model parameters θ and the GPR hyperparameters φ to be excessively adjusted in order to maximize agreement with observations. Although coupling a single physical constraint can partially alleviate this over-adjustment, it is generally insufficient to effectively constrain all parameters. When multiple physical constraints are simultaneously coupled, the posterior distributions of all parameters are confined to physically consistent ranges, thereby effectively mitigating parameter overfitting.

The GPR hyperparameters exhibit greater sensitivity to physical constraints because GPR lacks inherent physical regulation, whereas the parameters of the physical model (groundwater model) are already subject to internal constraints such as mass conservation. Consequently, the hyperparameters are more prone to excessive adjustment than the physical model parameters. The DDM coupled with multiple physical constraints primarily constrains the identification of GPR hyperparameters, leading to substantially more concentrated posterior distributions and effectively reducing parameter uncertainty.

Notably, the observational data available for parameter identification are limited in both case studies, with only 50 and 58 observations in case studies 1 and 2, respectively. Under limited data availability and structural bias in the physical model, parameter identification is subject to substantial uncertainty, resulting in pronounced equifinality. In this context, incorporating appropriate physical constraints effectively restricts the parameter identification process. Within the Bayesian uncertainty analysis, these physical constraints function analogously to additional observations, thereby further reducing parameter uncertainty.

The analysis of mass balance errors (Figs. 6 and 7) highlights the necessity of

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

coupling multiple physical constraints during structural error correction. In case study 1, when only the carbon mass conservation constraint is coupled, the mass balance error rate (ER) of chlorine remains at a relatively high level, whereas coupling both carbon and chlorine mass conservation constraints substantially reduces the ERs of both components. In case study 2, although the Cd mass conservation constraint reduces the ER for Cd, the ER decreases further after the concentration bound constraint is additionally coupled. These results indicate that a single physical constraint typically regulates only the predicted variables directly associated with it, whereas predictions of unconstrained variables may still violate physical principles. In contrast, coupling multiple physical constraints constrains the structural error correction process at the level of distinct physical mechanisms, thereby enhancing the physical consistency of DDM predictions.

The presence and the number of coupled physical constraints also significantly influence the forms of structural error (Figs. 12 and 13). Without physical constraints, structural errors exhibit large amplitudes and greater complexity, making them difficult for the GPR model to characterize accurately. Coupling a single physical constraint partially suppresses the amplitude of structural error variations. In contrast, when multiple physical constraints are coupled simultaneously, the posterior distributions of model parameters are further constrained, leading to substantially smaller and more stable structural error fluctuations. As a result, the complexity of structural error is markedly reduced, enabling the GPR model to predict structural errors more accurately.

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

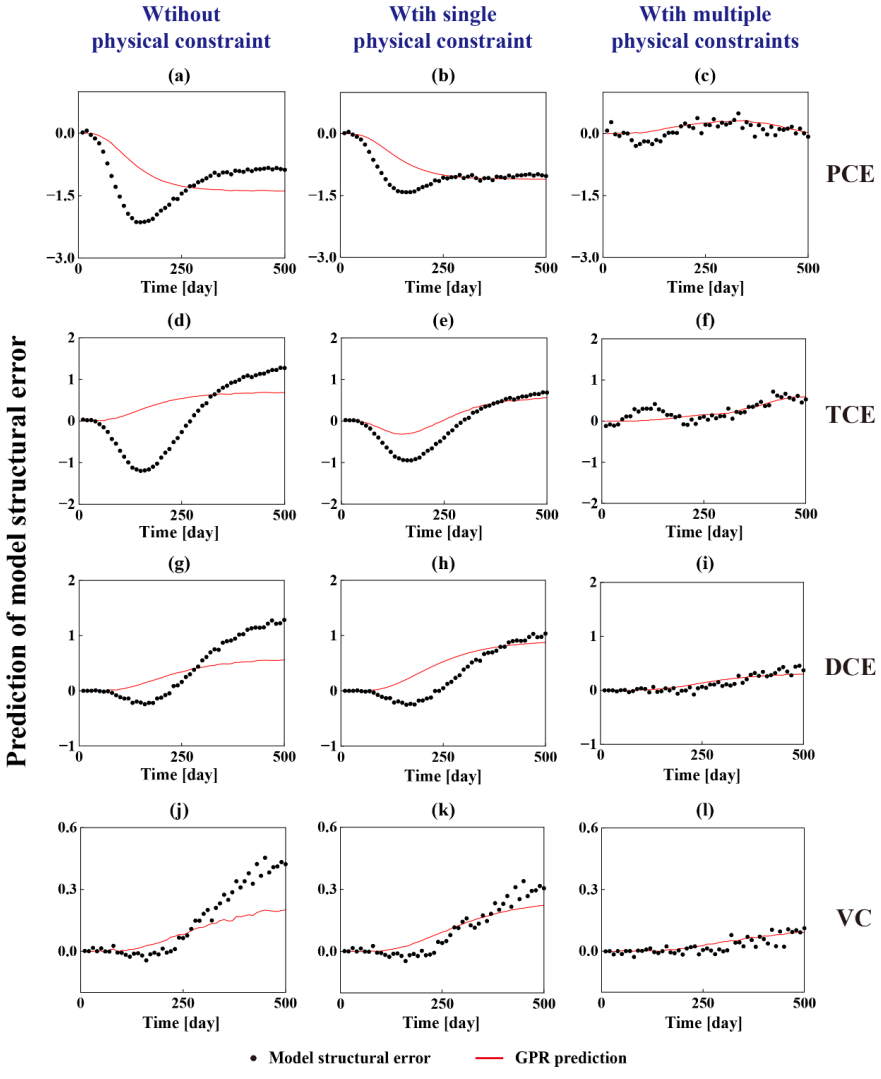

Figure 12. Predictions of model structural errors for the synthetic groundwater PCE reactive transport model without physical constraints (a, d, g, j), with single physical constraint (b, e, h, k) and with multiple physical constraints (e, f, i, l).

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

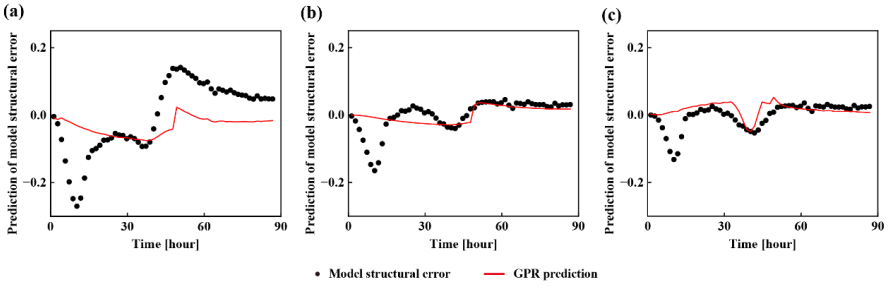

Figure 13. Predictions of model structural errors for the sand column Cd-phosphate cotransport model without physical constraints (a), with single physical constraint (b) and with multiple physical constraints (c).

5.4.3. Limitations of this Study

This study explicitly incorporates multiple physical constraints into the structural error correction process through a combined likelihood function, together with a dynamic sub-likelihood weight-updating strategy and associated stopping criterion, thereby overcoming the limitations of existing approaches in coupling multiple constraints. Unlike PIML approaches, the proposed method preserves the physical model predictions, with the corrected predictions consisting of the physical model, structural error, and measurement error, enhancing the robustness of extrapolative predictions. The method demonstrates broad applicability and enables the simultaneous coupling of constraints associated with distinct physical mechanisms, thereby improving the physical consistency of corrected predictions. Overall, this study provides a generalizable framework for enhancing predictive performance in simulations of contaminant reaction and transport within complex groundwater systems.

When the proposed method is applied to structural error correction in multicomponent reactive transport systems, multiple physical constraints typically need to be coupled simultaneously. In this context, the development of appropriate and effective physical constraints is essential. In general, the introduced constraints should exhibit clear mechanistic relevance to model structural error and observational data, and directly represent key physical processes that have been neglected or simplified. For instance, in more complex non-isothermal or multiphase flow systems, additional constraints may be needed to explicitly represent temperature-dependent transport and

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

reaction processes, or the mass exchange and conservation of components across different phases. The constraints should also possess explicit physical meaning and quantifiable formulations to avoid imposing inappropriate restrictions. Furthermore, when multiple constraints are incorporated, their complementarity and potential competition should be carefully evaluated to prevent redundant or conflicting constraints from adversely influencing parameter identification and compromising the physical consistency of predictions.

In addition, an increase in the number of constraints inevitably increases the complexity of the combined likelihood function, resulting in higher computational cost and potentially affecting MCMC convergence. Future research may therefore explore approaches to internally quantify and correct structural errors within groundwater models, thereby reducing reliance on externally imposed physical constraints and improving the applicability and stability of structural error correction.

6. Conclusions

Using DDM to correct model structural errors provides an effective approach for improving the predictive performance of groundwater contaminant transport models. This study proposes a general structural error correction framework capable of coupling multiple physical constraints. By constructing a combined likelihood function and dynamically updating sub-likelihood weights, the method explicitly incorporates multiple physical constraints into the structural error correction process, thereby substantially enhancing predictive accuracy and physical consistency of corrected predictions. The framework is demonstrated through two case studies: a synthetic three-dimensional PCE reactive transport simulation and a cadmium-phosphate cotransport sand column experiment. The main conclusions are summarized as follows:

(1) An innovative data-driven framework is developed that couples multiple physical constraints through a combined likelihood function and dynamic updating of sub-likelihood weights. The structural error correction method exhibits broad applicability to complex multicomponent reactive transport systems.

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

(2) The proposed dynamic weight-updating strategy with a stopping criterion maintains a balance between multiple physical constraints and observational information, thereby improving the robustness of parameter identification and constraint enforcement. Consequently, it enables the simultaneous coupling of distinct physical mechanism constraints.

(3) The DDM coupled with multiple physical constraints effectively regulates parameter identification during structural error correction and suppresses excessive parameter adjustment. Consequently, the forms of structural errors in calibrated groundwater models become less complex and easier to characterize, leading to improved predictive accuracy and enhanced physical consistency of the corrected predictions.

(4) Within the DDM framework for structural error correction, introducing appropriate physical constraints functions analogously to incorporating additional observations, thereby further reducing predictive uncertainty.

Acknowledgements

This study was supported by the National Key Research and Development Program of China (2024YFC3713001), and the National Natural Science Foundation of China (42477082, 42402236). We are grateful to the High-Performance Computing Center (HPCC) of Nanjing University for performing the simulations in this paper.

Data Availability

Version 5.0 of HYDRUS used for simulating the cadmium transport in the laboratory sand column cadmium-phosphate cotransport case is preserved at https://www.pc-progress.com/en/Default.aspx?hydrus. (Šimůnek, 2022)

Version 1.12.00 of MODFLOW-2005 used for simulating the groundwater flow in the synthetic three-dimensional tetrachloroethylene reactive transport case is preserved at http://dx.doi.org/10.5066/F7RF5S7G. (Harbaugh, 2017)

Version 1.1.00 of MT3D-USGS used for simulating the contaminant transport in

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

the synthetic three-dimensional tetrachloroethylene reactive transport case is preserved at http://dx.doi.org/10.3133/tm6A53. (Bedekar, 2016)

Data used for the two case studies are publicly available at, https://www.hydroshare.org/resource/60f2db9c87914233b7fac0dcbe1f4f7b/. (Tian, 2026)

References

Abu Doush, I., El-Abd, M., Hammouri, A., & Bataineh, M. (2023). The effect of different stopping criteria on multi-objective optimization algorithms. *Neural Computing & Applications*, *35*(2), 1125–1155. https://doi.org/10.1007/s00521-021-05805-1

Ayensa-Jiménez, J., Doweidar, M., Sanz-Herrera, J., & Doblaré, M. (2018). A new reliability-based data-driven approach for noisy experimental data with physical constraints. *Computer Methods in Applied Mechanics and Engineering*, *328*, 752–774. https://doi.org/10.1016/j.cma.2017.08.027

Bedekar, V., Morway, E.D., Langevin, C.D., and Tonkin, M. (2016). MT3D-USGS version 1.0.0: Groundwater Solute Transport Simulator for MODFLOW: U.S. [Software]. Geological Survey Software Release, 30 September 2016, http://dx.doi.org/10.5066/F75T3HKD

Cao, M., & Weng, Q. (2024). Embedded physical constraints in machine learning to enhance vegetation phenology prediction. *Giscience & Remote Sensing*, *61*(1). https://doi.org/10.1080/15481603.2024.2426598

Cooper, C., Zhang, J., & Gao, R. (2023). Error homogenization in physics-informed neural networks for modeling in manufacturing. *Journal of Manufacturing Systems*, *71*, 298–308. https://doi.org/10.1016/j.jmsy.2023.09.013

Engdahl, N. (2025). Impacts of uncertain permeability fields on the transient hydrologic response in coupled surface-subsurface simulations of a headwaters catchment. *Water Resources Research*, *61*(10). https://doi.org/10.1029/2025WR040668

Guger, M., Enzinger, C., Leutmezer, F., Kraus, J., Kalcher, S., Kvas, E., et al. (2020). Oral therapies for treatment of relapsing-remitting multiple sclerosis in Austria: a 2-year comparison using an inverse probability weighting method. *Journal of Neurology*, *267*(7), 2090–2100. https://doi.org/10.1007/s00415-020-09811-6

Guo, S., Agarwal, M., Cooper, C., Tian, Q., Gao, R., Grace, W., & Guo, Y. (2022). Machine learning for metal additive manufacturing: Towards a physics-informed data-driven paradigm. *Journal of Manufacturing Systems*, *62*, 145–163. https://doi.org/10.1016/j.jmsy.2021.11.003

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

Gupta, H., Clark, M., Vrugt, J., Abramowitz, G., & Ye, M. (2012). Towards a comprehensive assessment of model structural adequacy. *Water Resources Research*, *48*. https://doi.org/10.1029/2011WR011044

Harbaugh, A.W., Langevin, C.D., Hughes, J.D., Niswonger, R.N., and Konikow, L. F. (2017). MODFLOW-2005 version 1.12.00, the U.S. [Software]. Geological Survey modular groundwater model: U.S. Geological Survey Software Release, 03 February 2017. http://dx.doi.org/10.5066/F7RF5S7G

He, Y., & Xiu, D. (2016). Numerical strategy for model correction using physical constraints. *Journal of Computational Physics*, *313*, 617–634. https://doi.org/10.1016/j.jcp.2016.02.054

Hendrikx, T., Bierkens, M., de Louw, P., & Essink, G. (2026). Optimizing scavenger well strategies under parameter uncertainty to maximize allowable freshwater pumping rates in a coastal aquifer, The Netherlands. *Journal of Hydrology*, *664*. https://doi.org/10.1016/j.jhydrol.2025.134554

Hu, X., Shi, L., & Lin, G. (2021). The data-driven solution of energy imbalance-induced structural error in evapotranspiration models. *Journal of Hydrology*, *597*. https://doi.org/10.1016/j.jhydrol.2021.126205

Kadhar, K., & Baskar, S. (2018). A stopping criterion for decomposition-based multi-objective evolutionary algorithms. *Soft Computing*, *22*(1), 253–272. https://doi.org/10.1007/s00500-016-2331-7

Kasiviswanathan, K., He, J., Tay, J., & Sudheer, K. (2019). Enhancement of Model Reliability by Integrating Prediction Interval Optimization into Hydrogeological Modeling. *Water Resources Management*, *33*(1), 229–243. https://doi.org/10.1007/s11269-018-2099-x

Kennedy, M., & O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, *63*, 425–450. https://doi.org/10.1111/1467-9868.00294

Kim, T., Han, W., Shin, J., Yoon, S., Kang, P., Yang, W., & Lee, S. (2026). Quantitative evaluation of hydrogeologic properties controlling field-scale DNAPL contamination across porous and fractured media. *Journal of Hydrology*, *666*. https://doi.org/10.1016/j.jhydrol.2025.134819

Li, F., Morgan, K., & Zaslavsky, A. (2018). Balancing Covariates via Propensity Score Weighting. *Journal Of The American Statistical Association*, *113*(521), 390–400. https://doi.org/10.1080/01621459.2016.1260466

Lu, X., Liao, W., Zhang, Y., & Huang, Y. (2022). Intelligent structural design of shear wall residence using physics-enhanced generative adversarial networks. *Earthquake Engineering & Structural Dynamics*, *51*(7), 1657–1676. https://doi.org/10.1002/eqe.3632

Melching, D., Schultheis, E., & Breitbarth, E. (2023). Generating artificial

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

displacement data of cracked specimen using physics-guided adversarial networks. *Machine Learning-Science and Technology*, *4*(4). https://doi.org/10.1088/2632-2153/ad15b2

Mendenhall, C., Hardan, J., Chiang, T., Blumenschein, L., Tepole, A., & IEEE. (2024). Physics-Informed Neural Network for Scalable Soft Multi-Actuator Systems (pp. 716–721). Presented at the 2024 IEEE 7th International Conference on Soft Robotics, Robosoft. https://doi.org/10.1109/ROBOSOFT60065.2024.10522053

Neuman, S. P., & Wierenga, P. J. (2003). A comprehensive strategy of hydrogeologic modeling and uncertainty analysis for nuclear facilities a. Retrieved from https://api.semanticscholar.org/CorpusID:35968778

Nkedikizza, P., Biggar, J., Selim, H., Vangenuchten, M., Wierenga, P., Davidson, J., & Nielsen, D. (1984). On The Equivalence Of 2 Conceptual Models For Describing Ion-Exchange During Transport Through An Aggregated Oxisol. *Water Resources Research*, *20*(8), 1123–1130. https://doi.org/10.1029/WR020i008p01123

Pan, Y, Zeng, X., Xu, H., Sun, Y., Wang, D., & Wu, J. (2023). Use of stacked Gaussian processes regression method to improve prediction of groundwater solute transport model. *Journal of Hydrology*, *620*. https://doi.org/10.1016/j.jhydrol.2023.129530

Pan, Yue, Zeng, X., Xu, H., Sun, Y., Wang, D., & Wu, J. (2020). Assessing human health risk of groundwater DNAPL contamination by quantifying the model structure uncertainty. *Journal of Hydrology*, *584*, 124690. https://doi.org/10.1016/j.jhydrol.2020.124690

Pathiraja, S., Moradkhani, H., Marshall, L., Sharma, A., & Geenens, G. (2018). Data-Driven Model Uncertainty Estimation in Hydrologic Data Assimilation. *Water Resources Research*, *54*(2), 1252–1280. https://doi.org/10.1002/2018WR022627

Rasmussen, C. E., & Williams, C. K. I. (Eds.). (2005). Covariance functions. In *Gaussian Processes for Machine Learning* (p. 0). The MIT Press. https://doi.org/10.7551/mitpress/3206.003.0007

Refsgaard, J. C., Van Der Sluijs, J. P., Brown, J., & Van Der Keur, P. (2006). A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources*, *29*(11), 1586–1597. https://doi.org/10.1016/j.advwatres.2005.11.013

Saha, P., Dash, S., & Mukhopadhyay, S. (2021). Physics-incorporated convolutional recurrent neural networks for source identification and forecasting of dynamical systems. *Neural Networks*, *144*, 359–371. https://doi.org/10.1016/j.neunet.2021.08.033

Šimůnek, J., M. Šejna, G. Brunetti, and M. Th. van Genuchten. (2022). The HYDRUS Software Package for Simulating the One-, Two, and Three-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably Saturated Media. [Software]. Technical Manual I, Hydrus 1D, Version 5.0, PC Progress, Prague,

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

Czech Republic, 334p., 2022. https://www.pc-progress.com/en/Default.aspx?hydrus

Sun, X., Zeng, X., Wu, J., & Wang, D. (2021). A Two-Stage Bayesian Data-Driven Method to Improve Model Prediction. *Water Resources Research*, *57*(12). https://doi.org/10.1029/2021WR030436

Tian, J., Zeng, X., Wang, D., & Wu, J. (2026). A data-driven approach coupled with physical constraints to improve groundwater models with structural error. *Water Resources Research*, 62, e2025WR040247. https://doi.org/10.1029/2025WR040247

Tian, J. (2026). Data set for "A data-driven method coupling multiple physical constraints for correcting structural errors in groundwater contaminant transport models", HydroShare, http://www.hydroshare.org/resource/60f2db9c87914233b7fac0dcbe1f4f7b

Vangenuchten, M. (1980). A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils. *Soil Science Society of America Journal*, *44*(5), 892–898. https://doi.org/10.2136/sssaj1980.03615995004400050002x

Watson, T., Doherty, J., & Christensen, S. (2013). Parameter and predictive outcomes of model simplification. *Water Resources Research*, *49*(7), 3952–3977. https://doi.org/10.1002/wrcr.20145

Wen, Z., Zhu, Q., Krause, S., & Chen, W. (2025). Advancing Multi-Phase LNAPL Extraction Models by Incorporating a Wellbore Module. *Water Resources Research*, *61*(11). https://doi.org/10.1029/2025WR041014

Wu, Jiacheng, Wang, J.-X., Shadden, S. C., & AAAI. (2019). Adding Constraints to Bayesian Inverse Problems (pp. 1666–1673). Presented at the Thirty-Third AAAI Conference on Artificial Intelligence / Thirty-First Innovative Applications of Artificial Intelligence Conference / Ninth AAAI Symposium on Educational Advances in Artificial Intelligence.

Wu, JL, Levine, M., Schneider, T., & Stuart, A. (2024). Learning about structural errors in models of complex dynamical systems. *Journal of Computational Physics*, *513*. https://doi.org/10.1016/j.jcp.2024.113157

Wu, X., Gao, B., Lyu, X., Zeng, X., Wu, J., & Sun, Y. (2022). Insight into the mechanism of phosphate and cadmium co-transport in natural soils. *Journal of Hazardous Materials*, *435*. https://doi.org/10.1016/j.jhazmat.2022.129095

Xu, TF, Valocchi, A., Ye, M., & Liang, F. (2017). Quantifying model structural error: Efficient Bayesian calibration of a regional groundwater flow model using surrogates and a data-driven error model. *Water Resources Research*, *53*(5), 4084–4105. https://doi.org/10.1002/2016WR019831

Xu, Tianfang, & Valocchi, A. J. (2015). A Bayesian approach to improved calibration and prediction of groundwater models with structural error. *Water Resources*

https://doi.org/10.5194/egusphere-2026-3578
Preprint. Discussion started: 26 June 2026

Research, *51*(11), 9290–9311. https://doi.org/10.1002/2015WR017912

Ye, M., Neuman, S., & Meyer, P. (2004). Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff. *Water Resources Research*, *40*(5). https://doi.org/10.1029/2003WR002557

Yu, H., Wang, X., Ren, B., Zeng, T., Lv, M., & Wang, C. (2022). An efficient Bayesian inversion method for seepage parameters using a data-driven error model and an ensemble of surrogates considering the interactions between prediction performance indicators. *Journal of Hydrology*, *604*. https://doi.org/10.1016/j.jhydrol.2021.127235

Zeeshan, M., & Ruhl, A. (2023). Fates of potentially persistent and mobile organic substances in embedded outdoor columns for artificial groundwater recharge simulation. *Water Research*, *245*. https://doi.org/10.1016/j.watres.2023.120615

Zhang, J, Zhao, Y., Shone, F., Li, Z., Frangi, A., Xie, S., & Zhang, Z. (2023). Physics-Informed Deep Learning for Musculoskeletal Modeling: Predicting Muscle Forces and Joint Kinematics From Surface EMG. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *31*, 484–493. https://doi.org/10.1109/TNSRE.2022.3226860

Zhang, JJ, Liu, C., & Gao, R. (2022). Physics-guided Gaussian process for HVAC system performance prognosis. *Mechanical Systems and Signal Processing*, *179*. https://doi.org/10.1016/j.ymssp.2022.109336

Zhang, Q., Shi, L., Holzman, M., Ye, M., Wang, Y., Carmona, F., & Zha, Y. (2019). A dynamic data-driven method for dealing with model structural error in soil moisture data assimilation. *Advances in Water Resources*, *132*. https://doi.org/10.1016/j.advwatres.2019.103407

Zhang, X., Gong, J., & Xuan, F. (2021). A physics-informed neural network for creep-fatigue life prediction of components at elevated temperatures. *Engineering Fracture Mechanics*, *258*. https://doi.org/10.1016/j.engfracmech.2021.108130