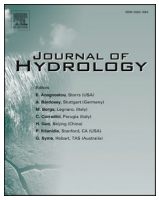
Journal of Hydrology 639 (2024) 131485

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol



Employing generative adversarial neural networks as surrogate model for reactive transport modeling in the hyporheic zone

Farzad Moeini [a,*], Reza Ershadnia [a], Rebecca L. Rubinstein [b], Roelof Versteeg [b], Pei Li [a], Jeffery T. McGarr [a], Alireza Meyal [b], Corey D. Wallace [c], Zhenxue Dai [d,e], Kenneth C. Carroll [f], Mohamad Reza Soltanian [a,g,*]

[a] Department of Geosciences, University of Cincinnati, Cincinnati, OH, USA
[b] Subsurface Insights, LLC, Hanover, NH, USA
[c] Geosyntec Consultants, Columbus, OH, USA
[d] School of Environmental and Municipal Engineering, Qingdao University of Technology, Qingdao, China
[e] College of Construction Engineering, Jilin University, Changchun, China
[f] Plant and Environmental Sciences Department, New Mexico State University, Las Cruces, NM, USA
[g] Department of Environmental Engineering, University of Cincinnati, Cincinnati, OH, USA

ARTICLE INFO

Keywords:
Generative adversarial networks
Reactive transport modeling
Deep learning
Hyporheic exchange
Heterogeneous aquifer

ABSTRACT

This study explores the application of advanced deep learning techniques, specifically conditional deep convolutional generative adversarial networks (cDC-GANs), to model and predict the complex dynamics of the hyporheic zone (HZ) driven by river stage fluctuations. The HZ is a critical transitional mixing zone between surface water and groundwater, where significant biogeochemical reactions occur. Traditional process-based reactive transport modeling methods are computationally intensive and require fine-scale parameterization. To overcome these challenges, we employ the cDC-GAN as a deep-learning-based surrogate model to predict zones of enhanced biogeochemical activity (i.e., hotspots) and reaction product concentrations in the HZ under varying bimodal sedimentary heterogeneity scenarios. The cDC-GAN model can efficiently capture nonlinear relationships between input data and output parameters without extensive parameterization, making it computationally less demanding. The model's generative capabilities allow for the creation of new data instances, enabling the exploration of diverse scenarios and interpolation between observed data points. The model uses maps of liquid saturation and solute concentrations at different time steps to generate corresponding maps of reaction rates and reaction product concentrations. Results demonstrate the model's qualitative and quantitative ability to capture complex relationships without relying on rigid assumptions about linearity or specific probability distribution functions. Notably, even amid variations in aquifer heterogeneity, the model consistently exhibits robust performance, validating its ability to adapt to dynamic geological settings. Our use of a simplified reaction scheme serves as a proof of concept for the cDC-GAN model's ability to simulate solute transport in the HZ, opening avenues for future applications including more complex reaction networks within HZ system such as nutrient cycling, organic matter degradation, redox reactions, and beyond.

1. Introduction

In the last two decades, there has been a growing interest in understanding processes within or adjacent to the hyporheic zone (HZ), a dynamic zone wherein surface water and groundwater mix and nutrients and solutes exchange and react (Boulton et al., 1998; Boano et al., 2014; Cardenas, 2015; Hester et al., 2017, 2021; Brunner et al., 2017; Krause et al., 2022). Spatial (Claret and Boulton, 2009; Irvine et al., 2015) and temporal (Gu et al., 2012) heterogeneity collaborate through mixing to give rise to localized point sources of reactivity, known as hotspots and hot moments (McClain et al., 2003; Bernhardt et al., 2017; Stegen et al., 2016; Stegen et al., 2018), which are essential for regulating nutrient cycling and attenuating contaminants (Lautz and Fanelli, 2008; Wallis et al., 2020). For instance, excess nitrate in rivers derived from anthropogenic sources (e.g., agricultural runoff, sewage discharges, industrial effluents) undergoes transformation and removal within the HZ,

* Corresponding authors at: Department of Geosciences, University of Cincinnati, Cincinnati, OH, USA.
E-mail addresses: moeinifd@mail.uc.edu (F. Moeini), soltanma@ucmail.uc.edu (M.R. Soltanian).

https://doi.org/10.1016/j.jhydrol.2024.131485

Available online 13 June 2024

which is crucial for maintaining water quality and reducing the risk of eutrophication (Bell, 1992; Anderson et al., 2002; Boano et al., 2010; Gu et al., 2012, Harvey et al., 2013). The HZ also improves water quality by immobilizing or transforming contaminants such as heavy metals and toxic elements (Birch et al., 2001; Palumbo-Roe et al., 2012). Effective strategies in water resource management need to consider these complex interactions within the HZ to ensure the sustainability of freshwater ecosystems (Lewandowski et al., 2020; Wu et al., 2020; Elshall et al., 2020; Scanlon et al., 2023).

Exchange dynamics within the HZ are influenced by the subsurface sedimentary heterogeneity, which controls spatial variability in physical (e.g., permeability (k)) and chemical (e.g., sorption properties, organic matter content) properties (Cardenas et al., 2004; Ryan and Boufadel, 2006; Soltanian et al., 2015a and b; Wallace et al., 2020, 2021). Such heterogeneity results in shorter residence times in higher-k sediments and vice versa (Western et al., 2001; Pryshlak et al., 2015; Yabusaki et al., 2017; Tonina et al., 2016). These complexities drive spatiotemporal changes in solute concentrations and reaction kinetics across multiple scales, making it challenging to characterize the exchange processes from field observations alone (Sawyer and Cardenas, 2009; Salehin et al., 2004). To address this issue, reactive transport models (RTMs) are commonly used to describe the coupled physical and biogeochemical processes within the HZ (Lautz and Siegel, 2006; Trauth et al., 2014; Li et al., 2017; Dwivedi et al., 2018b; Yang et al., 2019; Wallace et al., 2020; Li et al., 2020; Ng et al., 2020; Painter, 2021; Jan et al., 2021; Perzan et al., 2021; Behzadi et al., 2021; Singh et al., 2022). However, these process-based RTMs are often computationally intensive as they require fine-scale spatiotemporal discretization and detailed parameterization to accurately capture the spatial heterogeneity in sediment properties (e.g., Hunt et al., 2007; Dwivedi et al., 2018a). Moreover, multiple heterogeneity realizations and forward simulations are often required to evaluate RTM predictive uncertainty and perform sensitivity analysis (e.g., Ren et al., 2022; Frei et al., 2009; 2023), which adds to the computational burden.

Data-driven surrogate models are being increasingly developed to utilize large datasets alongside artificial intelligence (AI) techniques (such as machine learning (ML)) to facilitate efficient, accurate predictions and generate valuable insights into complex non-linear relationships (Russel and Norvig, 2010; Jordan and Mitchell, 2015). They efficiently characterize these relationships between data points without the extensive parameterization usually required for subsurface problems (Razavi et al., 2012). Traditional ML algorithms (e.g., support vector machines, decision trees, random forests, artificial neural networks) are used to predict hyporheic exchange dynamics based on various hydrogeological factors (Chang et al., 2014; Asher et al., 2015; Mori et al., 2019; Ren et al., 2021; Akhtar et al., 2022). However, the rigid structure and limited feature extraction capacity of these algorithms can be restrictive, especially for highly non-linear relationships.

With the advancement of AI, researchers are increasingly deploying new tools such as generative AI and deep learning (DL), including generative adversarial networks (GANs). GANs excel at capturing complex patterns and relationships, allowing them to generate new and innovative content beyond the training data (Goodfellow et al., 2014; Gonog and Zhou, 2019; Aggarwal et al., 2021; Brophy et al., 2023). Their ability to generate novel data (not simply repetitions or variations of the training data) distinguishes them from traditional ML algorithms. GAN-based models have been employed in subsurface related applications (e.g., energy exploration) to establish end-to-end mappings between high-dimensional input and output datasets, such as reconstructing solid-void structures of porous media (Mosser et al., 2017), characterizing complex heterogeneity models (Chan and Elsheikh, 2017), creating facies (i.e., bodies of sediments) models from hydrogeophysical data (Song et al., 2021; Zhan et al., 2022), predicting CO_2 plume migration in aquifers (Zhong et al., 2019), and generating accurate water saturation profiles in subsurface waterflooding projects (Zhong et al., 2020).

To date, there is a scarcity of literature exploring the extensive application of advanced DL methods (including GANs) in the context of HZ processes (Shen et al., 2021). Though process-based RTMs are routinely used within the HZ community, there is comparatively lower focus on the creation and validation of surrogate models. In this study, we employ a GAN model as a surrogate to predict the intricate dynamics of hotspots and reaction product concentrations within the HZ, driven by periodic fluctuations in river stage. This approach tackles the need for multiple RTM simulations usually conducted to handle prediction uncertainties, which can be demanding in terms of resources and time. Consequently, our method significantly contributes to filling a knowledge gap by offering a more resource-efficient solution for predicting the dynamics of the HZ. To demonstrate GAN capabilities, we explore the surrogate model performance under various degrees of sedimentation to examine how varying proportions of high-k facies volumes impact its effectiveness. In our study, we undertake both qualitative and quantitative evaluation by comparing the GAN results with those from process-based simulations. This comprehensive evaluation not only showcases the efficacy of our surrogate model in capturing the complexities of the HZ under various heterogeneity scenarios but also provides a practical basis for comparison with traditional RTM approaches.

2. Methodology

2.1. Sedimentary architecture representation

In this study, bimodal aquifer heterogeneities were simulated using a variety of both high-k and low-k facies proportions. The high-k facies is almost 20 times more permeable than the low-k facies. Increasing the content of high-k facies enhances the connectivity of high-k bodies, and thus, creating smoother fluid pathways and more efficient and dynamic exchange process in the HZ (Wallace et al., 2020). We used T-PROGS to apply a Markov chain approach in generating the heterogeneity models (Carle, 1999). These models were developed based on physical characteristics, including volume proportions and mean lengths of the facies. The geologic parameters associated with the high-k and low-k facies are presented in Table 1. We conducted the study using three specific scenarios to examine the effect of different k heterogeneity cases on model output. These scenarios were defined by varying the volume fraction of high-k and low-k facies as following (values of k for high-k and low-k

Table 1
Parameters used in the numerical simulation of groundwater flow and transport in the model.

Parameter	Value
high-k permeability	9.076×10^{-12} m^2
low-k permeability	4.538×10^{-13} m^2
high-k porosity	0.4
low-k porosity	0.55
high-k longitudinal dispersivity (α_L)	0.1 m
low-k longitudinal dispersivity (α_L)	0.1 m
high-k transverse dispersivity (α_T)	0.05 m
low-k transverse dispersivity (α_T)	0.01 m
high-k residual water saturation	0.069
low-k residual water saturation	0.26
high-k m (VGM empirical parameter)	0.5
low-k m (VGM empirical parameter)	0.5
molecular diffusion coefficient (D_m)	1.5×10^{-9} m^2/s
rate constant (K)	3.287×10^{-9} hr^{-1}
half saturation constant for solute A (K_A)	6×10^{-5} mol/kg
half saturation constant for solute B (K_B)	6×10^{-5} mol/kg
Initial Surface Water Solute Concentration	
Solute A	1×10^{-4} mol/kg
Solute B	1×10^{-4} mol/kg
Solute C	0.0
Initial Groundwater Solute Concentration	
Solute A	0.0
Solute B	0.0
Solute C	0.0

facies are presented in Table 1):

Scenario 1: 75 % high-*k*, 25 % low-*k* facies.

Scenario 2: 50 % high-*k*, 50 % low-*k* facies.

Scenario 3: 25 % high-*k*, 75 % low-*k* facies.

where scenario 1 represents the highest permeability field and scenario 3 represents the lowest permeability field. To evaluate the impact of these variations on surrogate model performance, the cDC-GAN model training was carried out independently for each scenario. 125 heterogeneity realizations were produced for each scenario through T-PROGS. The resulting *k* maps were then embedded within PFLOTRAN (as described in Section 2.2). To ensure effective model training and evaluation, the cDC-GAN model utilized a set of 100 unique *k* maps, representing 80% of the entire dataset. The remaining 20% of the data, referred to as the test set, were reserved to evaluate the performance of the trained model. This data split between training and testing is standard practice in machine learning, and enables the model to learn from a substantial portion of the data during training and then evaluate its ability to generalize to new, unseen instances. Fig. 1 visualizes the *k* field for one realization of each heterogeneity scenario along with the dimensions and boundary conditions applied in this study.

A sinusoidal function was utilized to model fluctuations in river stage, featuring a 0.5 m amplitude and a 12-hour period. The initial hydraulic head was determined as the mean value between the high and low river stages. To exclusively account for the interaction between the river and groundwater, we implemented no-flow boundaries with no solute exchange to the lower boundaries and lateral aspects of the subsurface region (Fig. 1). To prevent any influence on water table fluctuations, the right boundary was positioned at a significant distance (30 m) from the channel. The boundary separating the riverbed and vertical river bank functions as a seepage face (i.e., regions where water flow transitions between saturated and unsaturated conditions), incorporating a boundary for the mixed transport of solutes. The river concentration boundary condition was set to maintain a constant concentration equal to the initial surface water concentrations provided in Table 1 throughout the simulation period. This assumes that the river acts as a constant source of solutes to the hyporheic zone.

2.2. Numerical simulation

The numerical simulation approach involved the implementation of a two-dimensional model to simulate the water flow and reactive transport processes. To simulate variably saturated flow in porous domain, we employed Richards (1931) equation. Richards' equation is widely used in modeling groundwater-surface water interactions in coastal aquifers and riparian floodplains impacted by river fluctuations (e.g., Kirkland et al., 1992; Clement et al., 1994; Bause and Knabner, 2004; Suk and Park, 2019; Wallace et al., 2020; Jazayeri et al., 2021; Clément et al., 2023). Since our focus was on capturing the formation of hotspots where the reactions occur within the aqueous phase, and given that Richards' equation has been shown to be able to capture stage fluctuations within the unsaturated–saturated zone, we deemed it appropriate for our specific research objectives. Specifically, the equation takes the form:

$$\frac{\partial \theta}{\partial t} - \nabla . K \nabla h = \alpha C_m \frac{\partial \sigma_T}{\partial t} \tag{1}$$

where $\partial\theta/\partial t$ represents the change in water saturation, θ (m^3/m^3), over time (s), K indicates the hydraulic conductivity under variable saturation conditions (m/s), and h corresponds to the hydraulic head (m). σ_T represents the total stress resulting from river stage fluctuation (Pa), which is calculated as $\sigma_T = \rho g h$, where h represents the river stage height (m), ρ is the water density (kg/m^3), and g denotes the gravitational acceleration (m/s^2). C_m denotes the pore compressibility (Pa^{-1}), and α represents the loading efficiency (–). The variably saturated hydraulic conductivity, K, is described using the van Genuchten-Mualem (VGM) model (Van Genuchten, 1980):

$$K = \sqrt{S_e}\left(1 - \left(1 - S_e^{1/m}\right)^m\right)^2 \tag{2}$$

$$S_e = \frac{\theta - \theta_r}{1 - \theta_r} \tag{3}$$

where S_e is the effective saturation (m^3/m^3), θ_r is the residual water saturation (m^3/m^3), and m is an empirical parameter. The values used for these parameters in the model are presented in Table 1.

PFLOTRAN simulates the flow in porous media by utilizing the transient velocity field obtained through the solution of Richards' equation:

$$\frac{\partial C}{\partial t} = \nabla . (D \nabla C) - \nabla . \nu C + \sum R_i \tag{4}$$

where ν is fluid velocity (m/s) and C is solute concentration (mol/L). $\sum R_i$ represents the cumulative effects of every relevant reactions in the system. The dispersion coefficient, $D = D_{ij}$, is calculated using following formula:

$$D_{ij} = (\alpha_L - \alpha_T)\frac{\nu_i \nu_j}{\nu} + \alpha_T \nu \delta_{ij} + \tau D_m \tag{5}$$

where α_L and α_T represent longitudinal dispersivity (m) and transverse dispersivity (m), respectively. D_m denotes the diffusion coefficient (m^2/s), τ is tortuosity, and δ_{ij} stands for the Kronecker delta function (Bear, 1972).

To model the biogeochemical processes, we used a simplified representation of a microbial reaction (A + B → C). This reaction serves as a general representation of homogeneous chemical reactions (Gramling et al., 2002; Li and Berkowitz, 2019), which is effective for capturing the behavior of complex reactions described by multiple elementary reactions (Gillespie; 2000). This approach allows us to simulate reaction hotspots without dealing with the complexities of intricate reaction networks and generate more transformative findings. The reaction is expressed through multiple Monod equation:

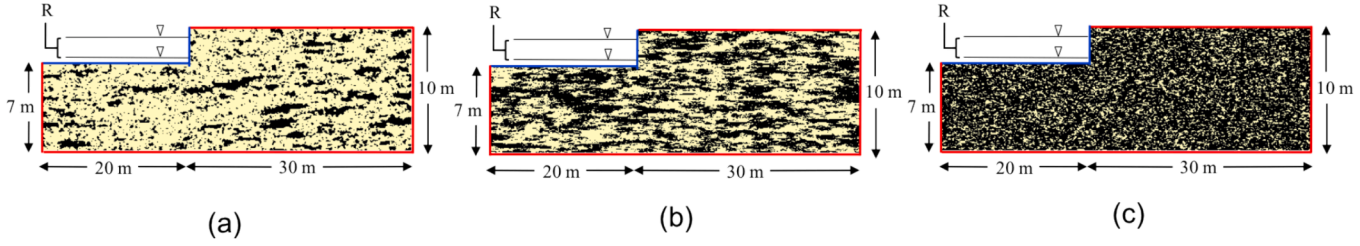


Fig. 1. The aquifer *k* field for one realization consisting of (a) scenario 1: 75% high-*k*, 25% low-*k*, (b) scenario 2: 50% high-*k*, 50% low-*k*, and (c) scenario 3: 25% high-*k*, 75% low-*k* fraction (the black color denotes low-*k* facies). The red lines represent the no-flow boundary condition, indicating areas where flow is restricted, and the blue lines indicate the boundary condition for seepage face, representing areas where water infiltrates. The high and low river stage lines depict the span of river stage fluctuations (R). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$R_i = K_r \frac{C_A}{K_A + C_A} \frac{C_B}{K_B + C_B} \tag{6}$$

where R_i denotes the reaction rate corresponding to the i^{th} reaction, C_A and C_B represent the concentrations of solutes A and B (mg/L) along with the respective half-saturation constants K_A and K_B (mg/L). K_r stands for the rate constant (hr^{-1}). An overview of the flow and transport parameters employed in this study is presented in Table 1.

To solve the groundwater flow field and transport equations and calculate fluxes of water and solute at the interface between sediment and water, we employed the finite-volume approach implemented in the PFLOTRAN software. The finite-volume method discretizes the domain into control volumes, allowing us to calculate the flow and transport properties at each volume. The model domain is evenly discretized with horizontal and vertical intervals of 0.05 (m) horizontal and vertical resolution. The subsurface domain is represented by a total of 176,000 computational grid cells. The solute transport simulations were performed over a duration of 1,080 h with a 240-hour spin-up period to ensure the attainment of a dynamic equilibrium for fluid flow in the groundwater simulations.

2.3. Surrogate model

The methodology employed in this study builds upon the concept of GANs for image-to-image translation processes. GANs are a class of deep learning models used for generating synthetic data that resembles a target dataset. GANs consist of two main components: a generative model (G) and a discriminative model (D), which compete against each other during the training process (Goodfellow et al., 2014). The G takes a random noise vector (z) as input, drawn from an initial distribution (typically a simple distribution like Gaussian), and produces generated samples ($G(z)$). The goal of G is to generate samples that closely resemble the true dataset distribution ($P_{data}(x)$). During training, G works to reduce the disparity between the distribution of the generated samples ($P_{G(z)}(x)$) and $P_{data}(x)$. This is achieved by updating the parameters of G to generate more realistic samples. D operates to discern between real samples taken from the actual dataset and the generated samples produced by G. It takes an input x (either real or generated) and outputs a probability $D(x)$ indicating the likelihood that x is a real sample. The ultimate goal of GANs is to optimize an objective function that involves both the generative and discriminative models. This objective function is formulated as:

$$\arg\min_G \max_D V(G, D) \tag{7}$$

$$V(G, D) = \mathbb{E}_{x \sim Pdata(x)} log[D(x)] + \mathbb{E}_{x \sim PG(z)(x)} log[1 - D(G(z))] \tag{8}$$

$V(G, D)$ represents the cost function that evaluates the quality of the generated samples. During training, D seeks to maximize $V(G, D)$ by accurately discriminating between real and generated samples, while G aims to minimize $V(G, D)$ by generating samples that resemble the true dataset. This adversarial process encourages G to generate increasingly realistic samples as D improves its ability to differentiate actual samples from those generated. By iteratively updating the parameters of G and D based on the objective function, GANs can learn the underlying data distribution and produce synthetic data that captures the statistical characteristics of the authentic data.

While GANs have shown remarkable success in generating realistic synthetic data, they also come with certain limitations. Standard GANs do not provide explicit control over the characteristics of the generated samples (Mirza and Osindero, 2014). Additionally, GANs may suffer from mode collapse, a situation in which the generator generates limited diversity and the generated samples fail to cover the full range of the target distribution (Zhang, 2021; Alqahtani et al., 2021). In this study, we employ a cDC-GAN framework to integrate the advantages of both conditional GANs (cGANs) and deep convolutional GANs (DC-GANs) architectures. The deep convolutional structure generates images from latent vectors using an encoder-decoder design to facilitate the learning of downsampling and upsampling operations by the generator and discriminator, leading to improved training stability for GANs (Radford et al., 2015). The cGAN extends the GAN architecture by introducing conditional information (y), represented as additional input variables, into both the generator $G(z|y)$ and discriminator $D(x|y)$ networks (Mirza and Osindero, 2014). This conditioning allows for the generation of images that exhibit desired characteristics or adhere to specific constraints, providing a flexible and powerful approach for achieving high-quality and controlled image synthesis. The incorporation of noise z under condition y in GAN loss function enables the generation of diverse and targeted outputs, which enriches the image-to-image translation process while aligning the generated distribution $P_{G(z|y)}(x)$ with the desired data distribution $P_{data}(x)$:

$$V(G, D) = \mathbb{E}_{x \sim Pdata(x)} log[D(x|y)] + \mathbb{E}_{x \sim PG(z|y)(x)} log[1 - D(G(z|y))] \tag{9}$$

Our cDC-GAN model adopts a similar architectural framework to the pix2pix model introduced by Isola et al. (2017). In our study, we focused on the uncertainty associated with the k field as the main input parameter. To address this, we used k maps as the input data (z vector) for the cDC-GAN model. Additionally, we incorporated conditioning data (y) consisting of maps of liquid saturation and solute concentrations for reactants A and B at different time steps, which are combined with the input to the generator model. This approach aimed to provide the cDC-GAN model with additional information to learn the dynamic relationships between the input k field and the resulting outputs of reaction rates and reaction product (i.e., solute C) concentrations. The target data (x) included maps of reaction rates and solute concentration at different times, which were concatenated into a single output. By combining these maps, the model can learn the temporal evolution of both output parameters simultaneously, capturing the dynamic relationships between them. This approach allows the cDC-GAN to generate predictions for reaction rates and reaction product concentrations in a single forward pass, eliminating the need for training distinct models for each output parameter and hence enhancing computational efficiency.

The generator in the cDC-GAN model consists of a U-Net encoder-decoder architecture with 3 sections including the downsampling process with 6 convolutional layers, along with 1 convolutional layer for the bottleneck (latent vector), and a corresponding number of deconvolutional layers in the expansion section to match the downsampling operations in the contraction section, resulting in a symmetrical architecture. Fig. 2 displays the architectural design of the cDC-GAN model. The architecture employs 4x4 kernels with a stride size of 2 in both the convolutional and deconvolutional layers, with the number of filters gradually increasing from 256 to 4,096 by a factor of two. In addition, skip connections, batch normalization, and leakyReLU activations with a leaky value of 0.02 are utilized to enhance information flow, normalization, and nonlinearity within the model, respectively. The final output of the generator is passed through a tanh activation function to ensure pixel values are in the range of [-1, 1]. The input and output images were resized to a dimension of 128 × 128 pixels prior to training the cDC-GANs. This choice ensures compatibility with the convolutional operations within the U-Net architecture of the generator in cDC-GAN model, which assume an equal spatial dimension for both input and output, preventing any inconsistencies in the network's computations. The discriminator network in the cDC-GAN model utilizes a PatchGAN approach, which consists of 3 convolutional layers followed by leakyReLU activations and a final sigmoid activation for distinguishing real images from generated ones at the patch level. By providing local-level discrimination, the PatchGAN discriminator encourages the generator to generate visually plausible outputs at fine-grained scales. The selection of other hyperparameters follows the approach outlined in the pix2pix model developed by Isola et al. (2017).

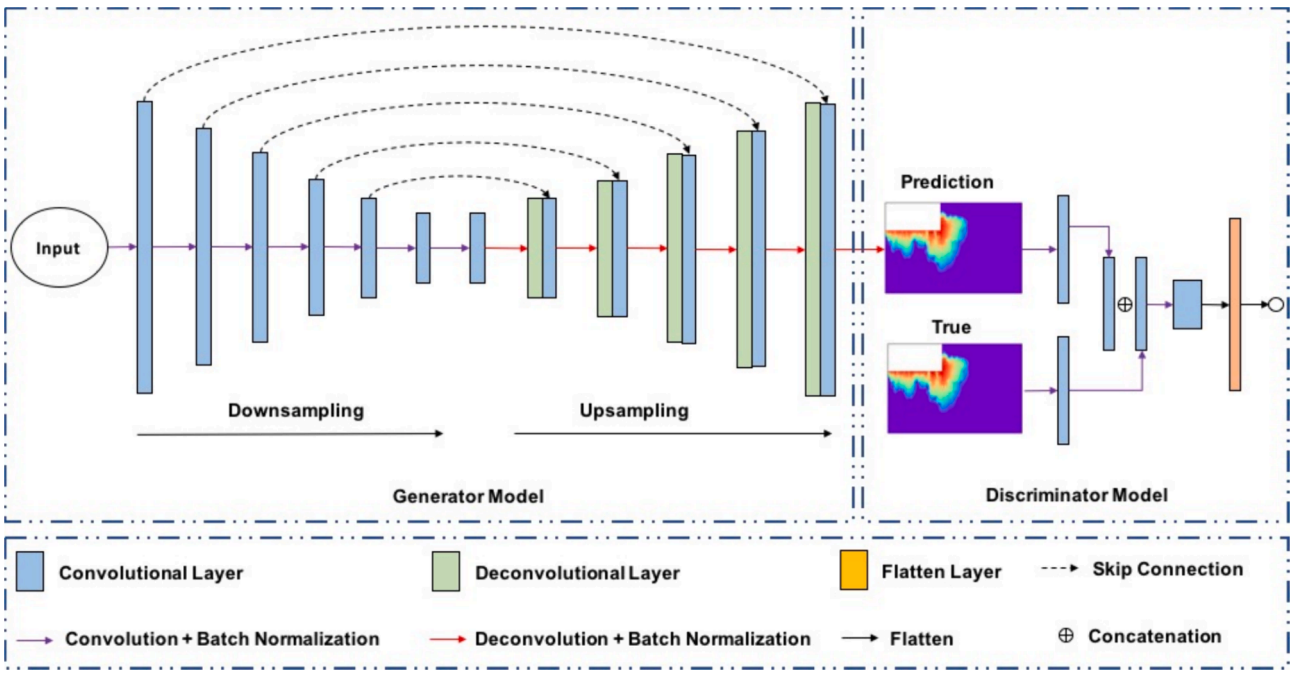


Fig. 2. The design of surrogate model (modified from Zhong et al., 2020). The cDC-GAN comprises a U-net architecture with skipped connections for the generator (on the left) and a PATCHGAN convolutional architecture for the discriminator (on the right). The U-net integrates downsampling and upsampling along with skipped connection paths to optimize its performance. The input data for the model include k field, liquid saturation, and reactants concentration maps at each time step, which are concatenated into a unified input. Similarly, the model outputs comprise predicted reaction rates and concentration maps, concatenated into a single output.

Prior to training, all the datasets are normalized and rescaled from their original range to a standard range between 0 and 1 to prevent issues such as vanishing or exploding gradients during training. For optimizing the parameters of the deep neural network, we selected the adaptive moment estimation method (ADAM), which is an extension of the stochastic gradient descent (SGD) algorithm. Through this methodology, the cDC-GAN model demonstrates its capability to generate high-fidelity and coherent images conditioned on specific input characteristics.

In this research, we utilized PyTorch, an open-source DL package, for the implementation of the cDC-GAN models. To improve training efficiency, the DistributedDataParallel (DDP) feature in PyTorch was employed to parallelize the training process across 4 graphic processing units (GPUs). Each GPU performed its own forward and backward pass on a subset of the training data. The gradients computed on each GPU were then synchronized and averaged to update the model parameters. All cDC-GAN models were trained using NVIDIA Volta V100 GPUs on a cluster node, completing a total of 400 epochs. The procedural steps for training and evaluating the implemented surrogate model in this study are depicted in Fig. 3.

2.4. Evaluation metrics

Structural Similarity Index Measure (SSIM) is commonly used in computer vision and image processing metric to assess the fidelity of generated images (Wang et al., 2004). SSIM considers the structural information of images and evaluates how closely the generated images align with the authentic dataset in terms of visual perception. SSIM quantifies the similarity in terms of texture, edges, and overall image structure. The SSIM between two windows, denoted as x (for real images) and y (for predicted images), is determined by the following formula:

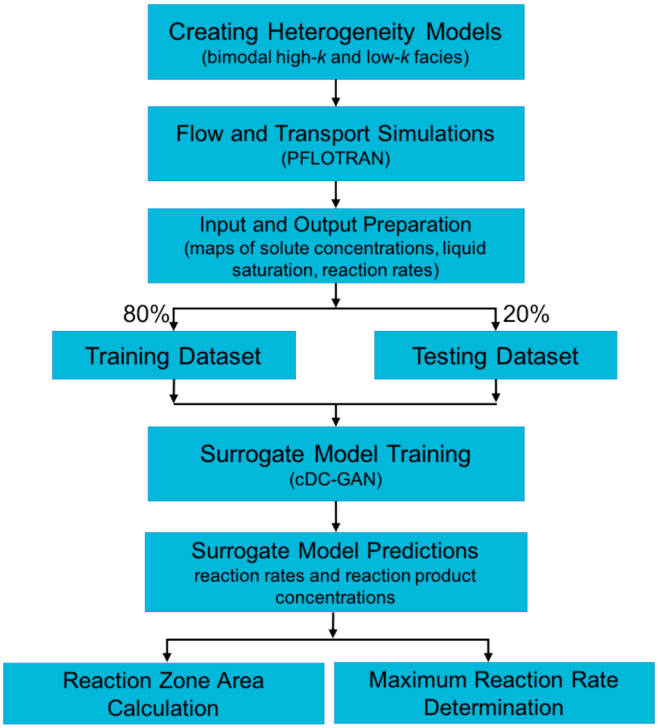


Fig. 3. The general workflow that integrates process-based simulation and data-driven surrogate model for predicting reaction rates and concentration maps in the HZ.

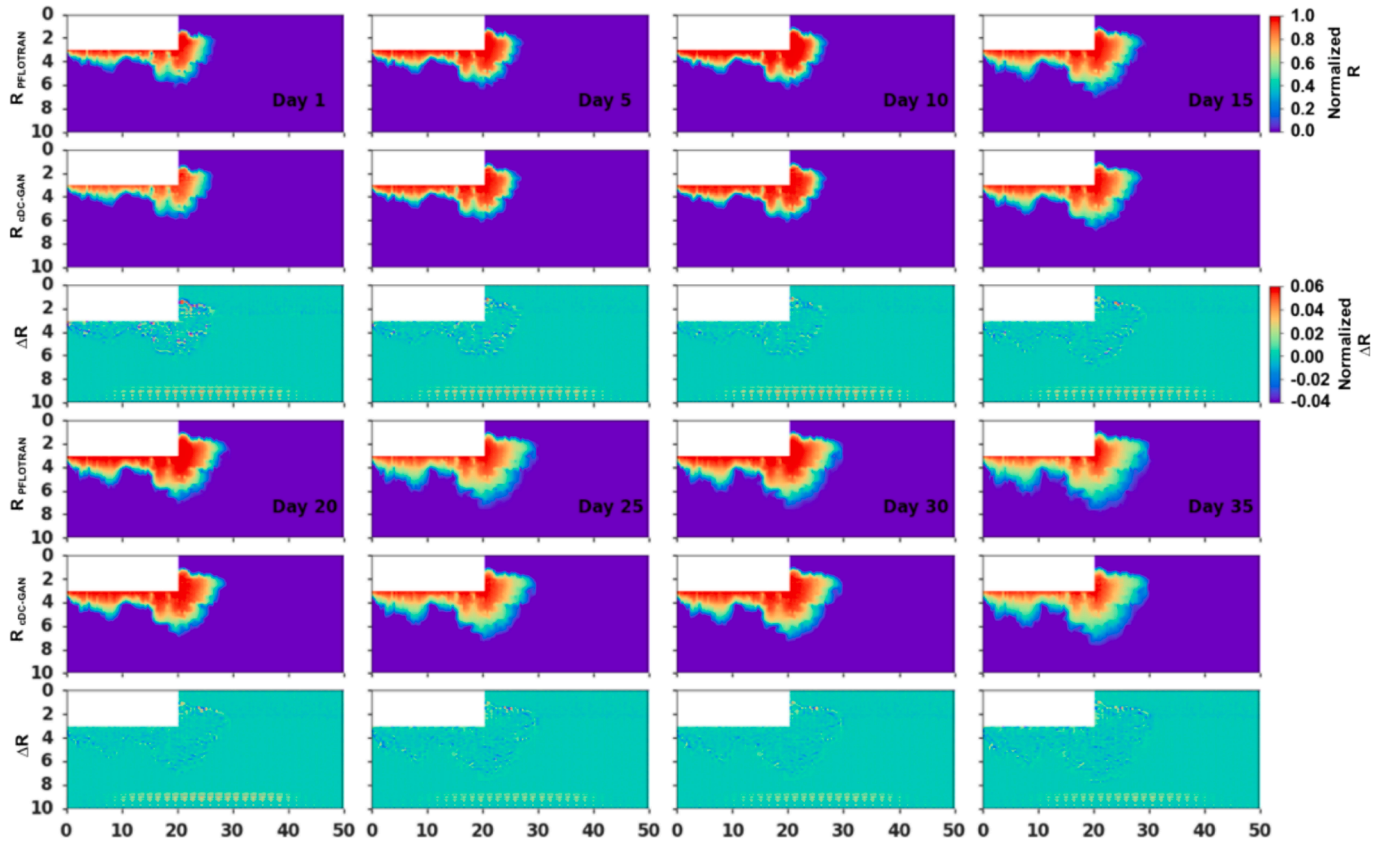


Fig. 4. Snapshots of normalized natural logarithm of reaction rate maps at different time steps for scenario 1. first row: PFLOTRAN simulation results; second row: cDC-GAN prediction results; third row: residual error between each pair of maps. Rows 4 to 6 introduce additional time steps for both models. Lengths are in meters, with a vertical exaggeration of 3.

$$\mathrm{SSIM}(x,y) = \frac{\left(2\mu_x\mu_y + c_1\right)\left(2\sigma_{xy} + c_2\right)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)} \tag{10}$$

where μ_x and μ_y refer to the means calculated for windows x and y, respectively. σ_x and σ_y correspond to the standard deviations of windows x and y, while σ_{xy} indicates the covariance between window x and window y. The constants $c_1 = 0.01$ and $c_2 = 0.03$ are used for numerical stability. The resulting SSIM score varies between 0 and 1, with a value nearing 1 signifying a greater resemblance between the generated image and the ground truth.

In addition, we used the root mean squared error (RMSE) and coefficient of determination (R^2) evaluation metrics for assessing the quantitative performance of the surrogate models. RMSE measures the average deviation of predicted values from ground truth data by taking the square root of the mean of the squared differences. The calculation of RMSE between two images is as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2} \tag{11}$$

where N represents the number of samples being compared. Variables x_i and $\hat{x}_i$ denote the true (reference) and predicted (generated by cDC-GAN) images, respectively. A lower RMSE value indicates a smaller average deviation and, therefore, a higher level of similarity with the ground truth data in the generated images. In the context of image generation, RMSE quantifies the pixel-level disparities between the generated image and the ground truth, providing a numerical value that represents the overall dissimilarity between the two.

R^2 is a metric that measures how well the independent variables in a regression model explain the variance in the dependent variable. It is expressed as following:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{\sum_{i=1}^{N}(x_i - \bar{x}_i)^2} \tag{12}$$

where the $\bar{x}_i$ represents the average value of reference images. R^2 score ranges between 0 and 1, where 0 indicates that the independent variables cannot explain any of the variability in the dependent variable, and 1 indicates a perfect fit, where the independent variables can explain all the variability.

3. Results

3.1. Interpretations

We demonstrate the functionality of cDC-GAN by using a single example from the testing set for each heterogeneity scenario. Examples of other realizations are provided in the supplementary information. Analysys of potential overfitting is provided in Text S1 in Supplementary materials. Figs. 5, 6, 7 present a comparative analysis of the reaction rate obtained from PFLOTRAN and the corresponding maps generated by cDC-GAN at various time steps. Results for reaction product C concentration are provided in and Figs. S13, S14, S15. The results of flow and transport simulations conducted using PFLOTRAN are provided in Figs. S4 to S12, including maps of solute A and solute B concentrations along with liquid saturation maps. The PFLOTRAN simulations show that in all heterogeneity scenarios, the extent of the reaction zone and reaction product C plume expand over time. The distribution of reaction rates exhibit periodic temporal variations attributed to solute infiltration across the riverbank driven by fluctuating river stages. The

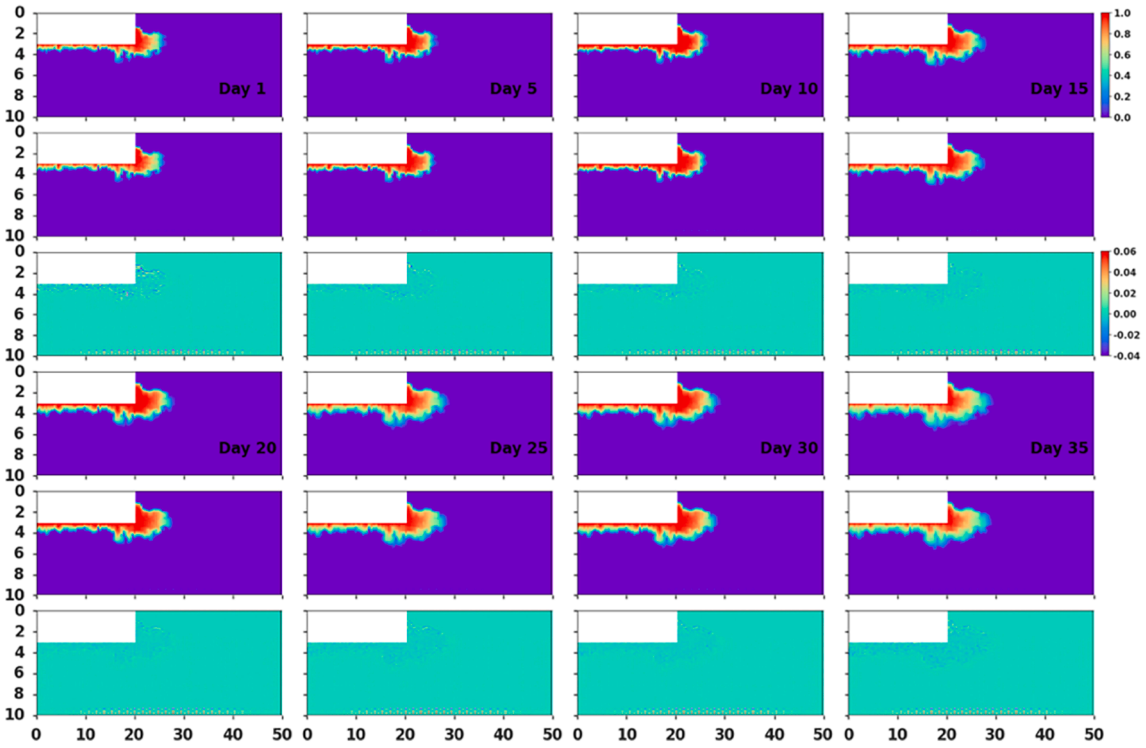


Fig. 5. Snapshots of normalized natural logarithm of reaction rate maps at different time steps for scenario 2. first row: PFLOTRAN simulation results; second row: cDC-GAN prediction results; third row: residual error between each pair of maps. Rows 4 to 6 introduce additional time steps for both models. Lengths are in meters, with a vertical exaggeration of 3.

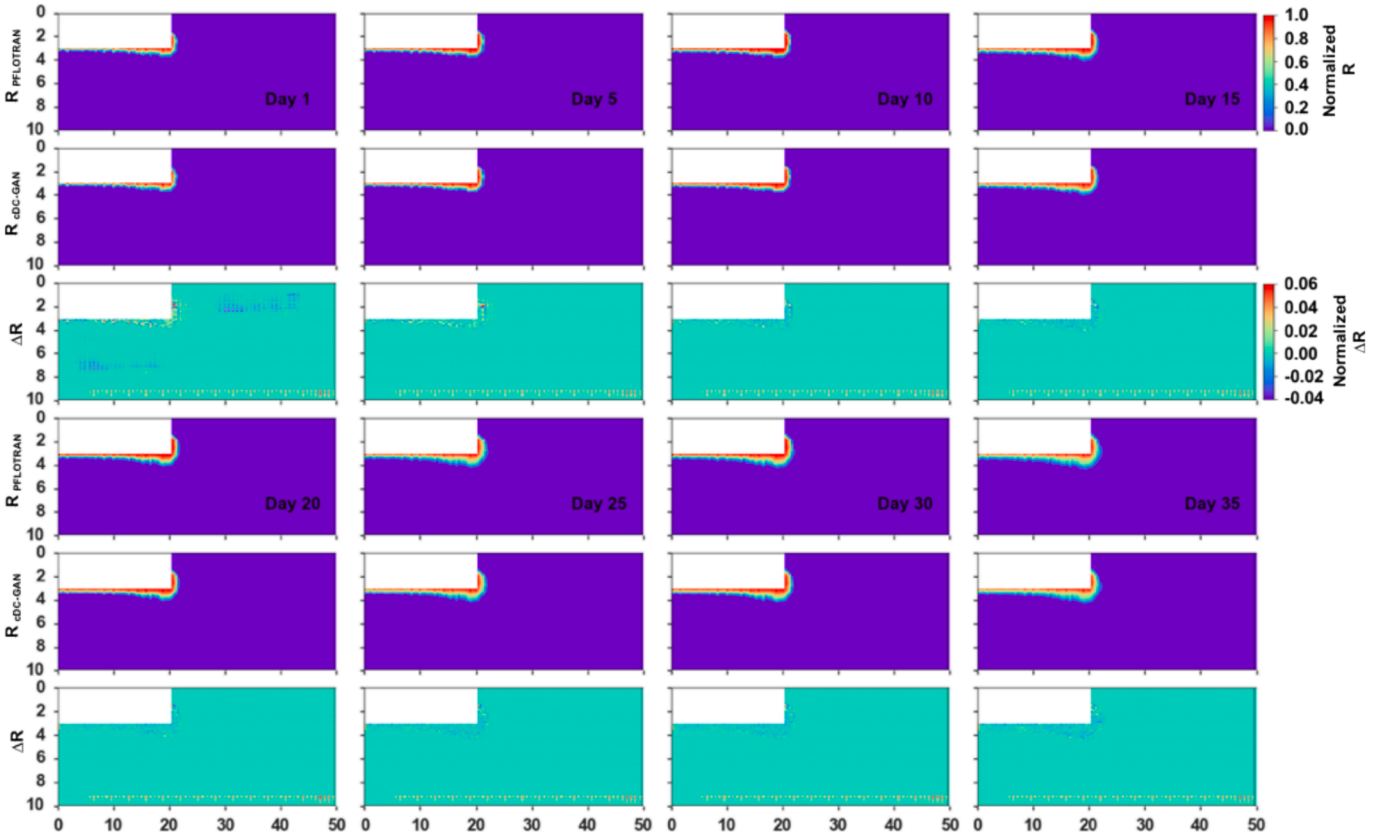


Fig. 6. Snapshots of normalized natural logarithm of reaction rate maps at different time steps for scenario 3. first row: PFLOTRAN simulation results; second row: cDC-GAN prediction results; third row: residual error between each pair of maps. Rows 4 to 6 introduce additional time steps for both models. Lengths are in meters, with a vertical exaggeration of 3.

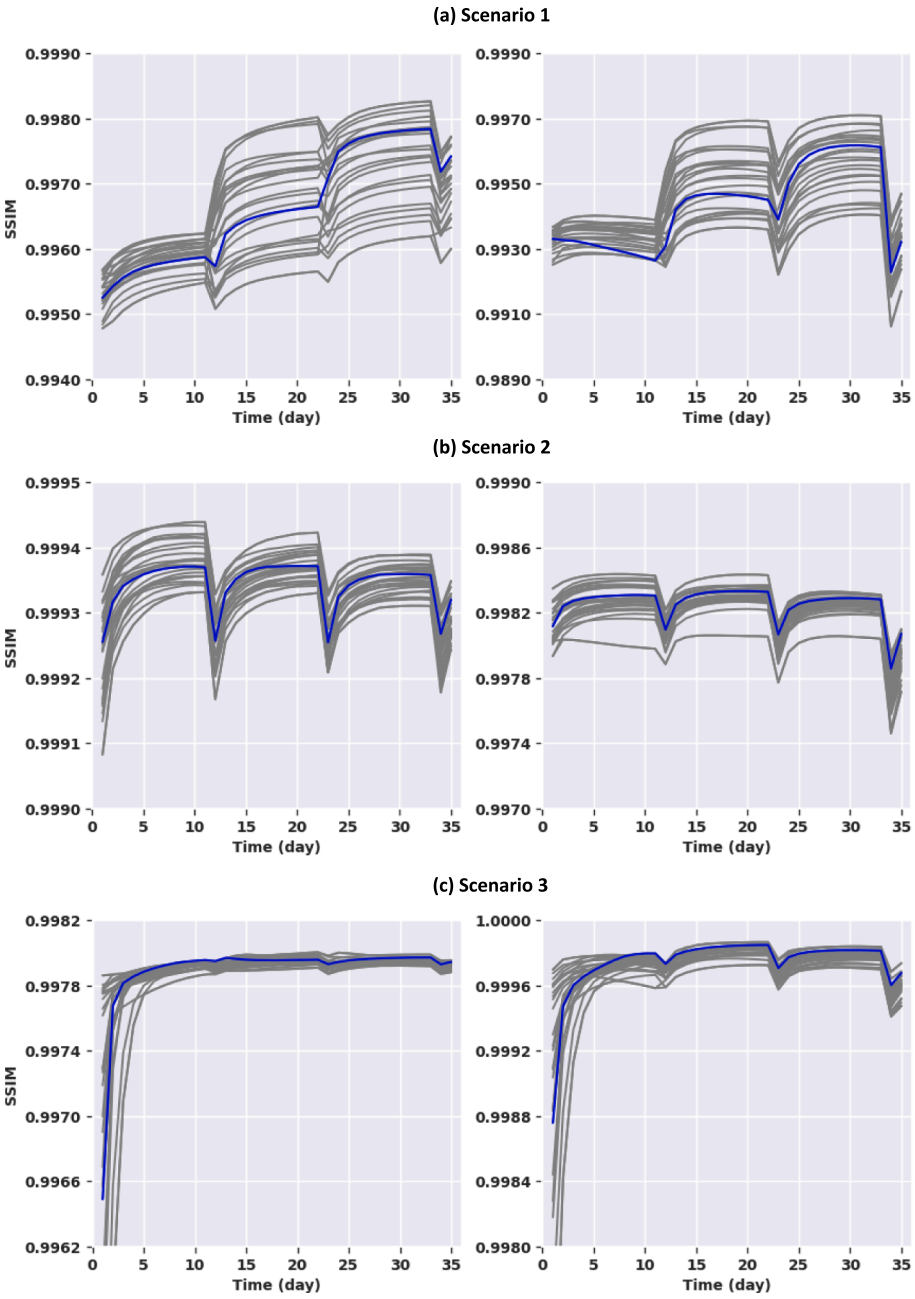


Fig. 7. SSIM values between maps obtained by PFLOTRAN and those predicted by cDC-GAN for the entire testing dataset at all simulated time steps. The left column presents SSIM values for reaction rate maps, while the right column shows SSIM values for reaction product concentration maps. The SSIM values are displayed for: (a) scenario 1, (b) scenario 2, and (c) scenario 3. In each subplot, the blue curve represents the SSIM values for the selected realization that is analyzed in detail for each heterogeneity scenario. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

hotspots, characterized by higher reaction rates compared to the surrounding areas, predominantly occur at the river-aquifer interface, particularly below the water table. The reaction product concentration, however, steadily increases over time due to continuous supply of reactants through water flow paths. In addition, as the ratio of high-*k* to low-*k* facies decreases from 75 % to 25 %, the reaction zone area contracts to a more limited spatial extent, leading to the generation of a more localized reaction rates and less dispersed solute C plume through the depositional system. This contraction is attributed to the reduced volume fraction high-*k* facies that result in a more limited interconnected permeable layers, which consequently confines and inhibits the transport and dispersion of solutes over larger distances.

Across all heterogeneity scenarios, the results indicate that cDC-GAN models perform reasonably well in approximating the distribution of the reaction rates and reaction product concentration at different timesteps. The residual error, representing the difference between true values and predicted values for both output parameters, consistently remains below 6 % across all cases (Figs. 4, 5, 6 and Figs. S13, S14, S15), demonstrating the model's reliability in replicating plume shapes and reaction rates. Moreover, the residual maps reveal that although the proposed cDC-GAN model exhibits proficiency in generating maps of output parameter distributions, it can also generate unwanted noise and localized artifacts, as observed in the residual error maps, which have a negative effect on the quality of generated maps (Isola et al., 2017; Ledig et al., 2017; Zhang et al., 2020; Zhong et al., 2020). This issue is more pronounced in situations where the relevant spatial information of plume sizes is confined to a limited area within the input size, making it difficult for the model to capture the intricate relationships between

neighboring points accurately. The model's receptive field, which determines the region of input that influences an output, may not encompass the necessary contextual information required to generate precise representations of the localized patterns and fine-scale details, leading to the presence of artifacts and noisy patterns. The highest residual error is consistently observed at earlier time steps across all heterogeneity scenarios, particularly near the river-aquifer interface which aligns with regions of higher hotspot occurrences. This phenomenon could be attributed to the cDC-GAN model's architecture and objective function, which may prioritize minimizing the overall error across the simulation period at the expense of accurately capturing specific details at earlier time steps when the plume size is more localized.

Fig. 7 shows the SSIM values between the predictions generated by the cDC-GAN models and the PFLOTRAN simulation results. These values were computed for the reaction rates and reaction product concentration maps at each time step, ranging from day 1 to day 35. These SSIM values encompass the entire set of 25 testing maps at all simulated time steps. The mean SSIM values at each time step for each scenario are presented in Fig 8. Collectively, these figures show that the cDC-GANs consistently achieve a minimum SSIM value above 0.97 across all testing realizations, indicating a high level of similarity between the predicted distributions of reaction rates and reaction product concentration by the cDC-GANs and the calculations performed by PFLOTRAN. Among the three heterogeneity scenarios, the highest mean SSIM was achieved in scenario 2 (Fig. 7b), which represents a balanced distribution of high-*k* and low-*k* facies. This suggests that in this scenario, the cDC-GAN can achieve a more optimal trade-off between capturing the details of the plume shape and avoiding the artifacts, resulting in a better balance between fidelity and noise reduction. On the other hand, in scenario 3 (Fig. 7c), which is characterized by a localized pattern of plume shape resulting from the dominance of low-*k* facies, the cDC-GAN model struggles to reproduce the contracted plume without introducing artifacts and inaccuracies, leading to lower mean SSIM values compared to the other scenarios. Furthermore, the cDC-GAN model demonstrates a better ability to capture the complex relationships associated with reaction rate distributions compared to reaction product concentration distributions, as evidenced by the consistently higher mean SSIM values for reaction rate maps across all heterogeneity scenarios (Figs. 7 and 8). Reaction rates are more directly influenced by the local hydrodynamics, solute mixing, and biogeochemical conditions within the hyporheic zone, which exhibit high spatial variability. The cDC-GAN model's convolutional architecture and ability to learn hierarchical features make it well-suited for capturing these intricate patterns and relationships governing reaction rate distributions. In contrast, reaction product concentrations are influenced by the integrated effects of reaction rates, flow paths, and residence times, leading to a smoothing effect on their spatial variations, making it more challenging for the model to capture the fine-scale details and local heterogeneity.

For quantitative analysis and assessing the agreement between cDC-GAN models and PFLOTRAN, we determine the volume of the reaction zone and maximum reaction rate at each time step. Fig. 9 displays the comparison between the simulated and predicted results. The first subplot presents the mean values and min–max uncertainty bounds computed for both cDC-GAN and PFLOTRAN. The second and third subplots show error bars and scatter plots, respectively, depicting the variation and correlation between the mean values.

Over time, the volume of the reaction zone in the hyporheic zone increases as the biogeochemical reaction progresses and reactants transform into product C, resulting in the expansion of the reaction zone. The periodic occurrences of localized hotspots within the hyporheic zone, driven by river stage fluctuation, makes sudden shifts in the reaction zone boundaries. This interplay between the biogeochemical reaction and river stage fluctuations gives rise to the observed periodicity in the reaction rate, with the maximum reaction rate experiencing peaks and troughs in response to varying river stage levels. Note that the relationship between the period of maximum reaction rates and the river stage fluctuations is nonlinear. The biogeochemical reactions within the hyporheic zone are influenced not only by the periodic influx of solutes from the river but also by various other factors, such as the heterogeneity in permeability fields, the transport and mixing processes, and the kinetics of the biogeochemical reactions. These factors contribute to the complex spatiotemporal dynamics of reaction rates, leading to temporal patterns that does not directly correspond to the period of the river stage fluctuations.

For the reaction zone, the predictions showed a strong correlation with actual values, as indicated by R^2 values exceeding 0.98 for all cases, along with accurate estimation demonstrated by the RMSE consistently below 0.2. Similarly, for maximum reaction rate, $R^2 > 0.95$ and RMSE $<$ 0.006 were achieved in all heterogeneity scenarios. The higher R^2 value for the volume of the reaction zone compared to the maximum reaction rate can be attributed to the cumulative nature of the volume measurement, which represents the overall spatial extent of the reaction over time. Nevertheless, the high scores of these performance metrics demonstrate the quantitative ability of cDC-GAN models to capture and model these dynamic aspects of the system with a relatively high accuracy. The generally lower maximum reaction rates predicted by the cDC-GAN model can be attributed to the challenges in accurately capturing localized hotspots and peak values. The model's training

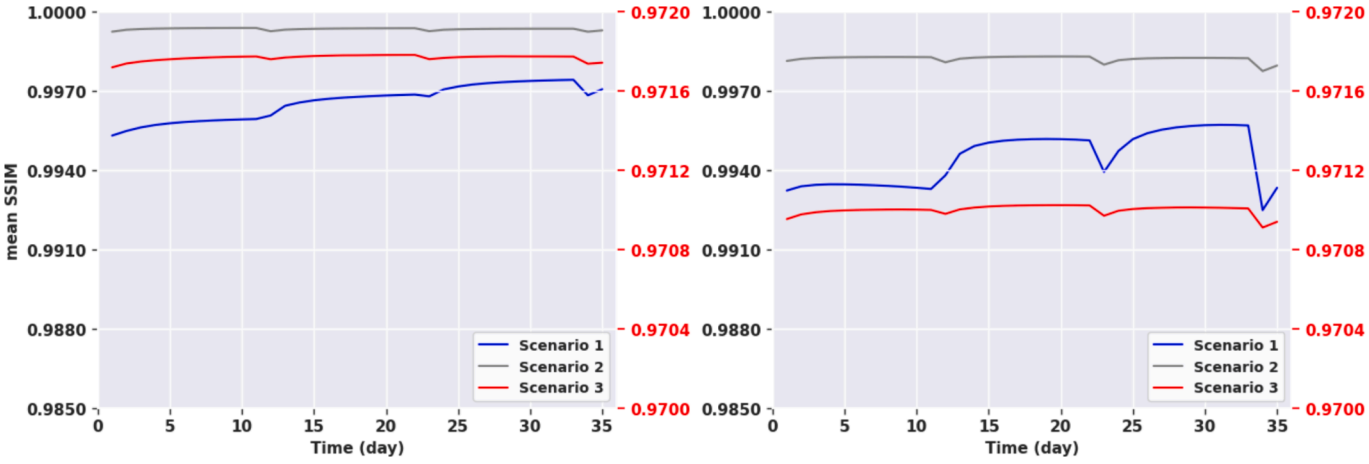


Fig. 8. Mean ssim values between maps obtained by pflotran and those predicted by cdc-gan calculated across the entire testing dataset for each heterogeneity scenarios at all simulated time steps (for scenario 3, the values are depicted in red on the right *y*-axis of each subplot). Left: mean SSIM values between reaction rate maps; Right: mean SSIM values between reaction product concentration maps. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

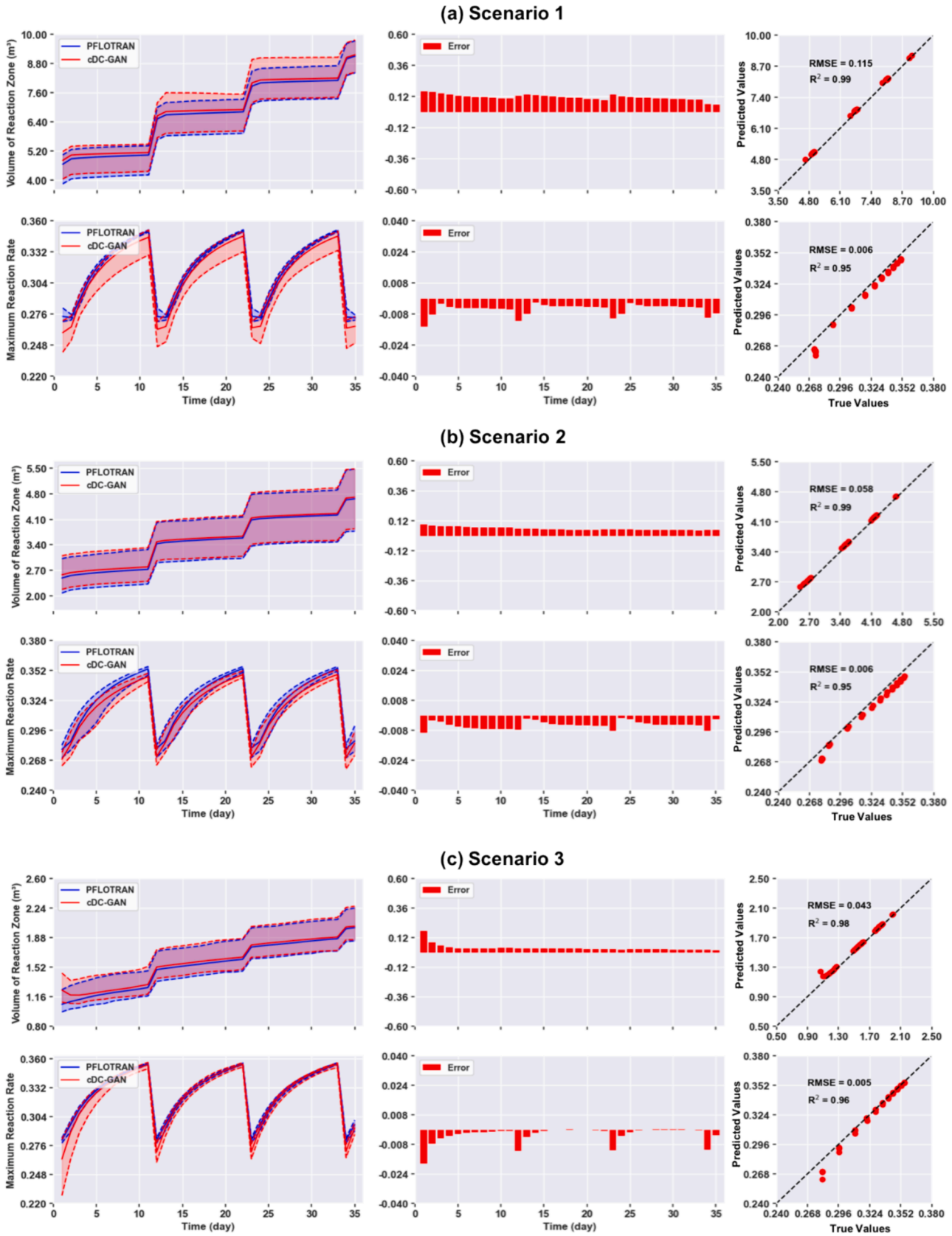


Fig. 9. Comparison of the calculated reaction zone area and the normalized natural logarithm of maximum reaction rate between the results obtained from PFLOTRAN and cDC-GAN results for: (a) scenario 1, (b) scenario 2, and (c) scenario 3. The first column illustrates the mean value along with the min–max uncertainty bounds at various time steps; The second and third columns depict the error bars (difference between PFLOTRAN results and cDC-GAN predictions) and scatter plot illustrating the average of simulated and predicted results.

focuses on minimizing overall error, potentially leading to a trade-off in predicting localized peak values. Also, the model's architecture and hyperparameters may not be optimized for capturing fine-scale variations and sharp gradients, and the smoothing effects of convolutional operations and regularization techniques could contribute to the attenuation of peak values. Exploring alternative loss functions or architectural modifications that emphasize reproducing maximum reaction rates could potentially improve the model's performance.

To gain a better understanding of the cDC-GAN model's internal workings, we also analyzed the feature maps obtained after the first convolutional layer, which consists of 64 × 64 feature maps. Fig. S16 presents the feature maps calculated at two different time steps, 10 days and 35 days, along with the residual feature map obtained by subtracting the feature maps at 10 days from those at 35 days. The feature maps exhibit spatial patterns that closely resemble those found in the original conditioning data, such as the permeability, liquid saturation, and reactant concentration maps. This similarity indicates that the convolutional layer effectively learns to extract and preserve the relevant spatial information from the input data. The residual map provides insights into the temporal evolution of the learned features. The presence of distinct patterns in the residual map indicates that the cDC-GAN model is capturing the changes in the distribution of the relevant variables over time. The residual map highlights the regions where the liquid saturation and reactant concentration maps undergo notable changes between the two time steps. This observation suggests that these variables control the overall temporal evolution of the reactive transport processes in the hyporheic zone. The feature maps corresponding to the permeability field remain largely unchanged between the two time steps, as the permeability distribution is a static property of the porous medium and does not vary over time.

3.2. Computational cost

Fig. 10 presents a comparison between the two different computational methods, PFLOTRAN and cDC-GAN. We conducted PFLOTRAN simulations using 12 CPU cores, taking approximately 40 min to finish each individual simulation. Executing 100 simulations with PFLOTRAN, which requires 67 h, is necessary to generate the training dataset for training the cDC-GAN model, as indicated by the point (100, 67) in Fig. 10. Training each cDC-GAN model on 4 GPU cores takes approximately 6 h, which corresponds to the equivalent of 9 numerical simulations, as illustrated by the data point (109, 73). Once trained, these cDC-GAN models can swiftly predict the desired parameters for various time instances and heterogeneity scenarios. In contrast, PFLOTRAN still requires roughly 40 min for each simulation. Thus, in scenarios where the requirement exceeds 109 simulations, opting for the cDC-GAN model proves to be a computational advantage. This efficiency arises from the substantial reduction in prediction time made possible by the cDC-GAN's efficient training, as opposed to PFLOTRAN, which maintains a consistent simulation time for each iteration.

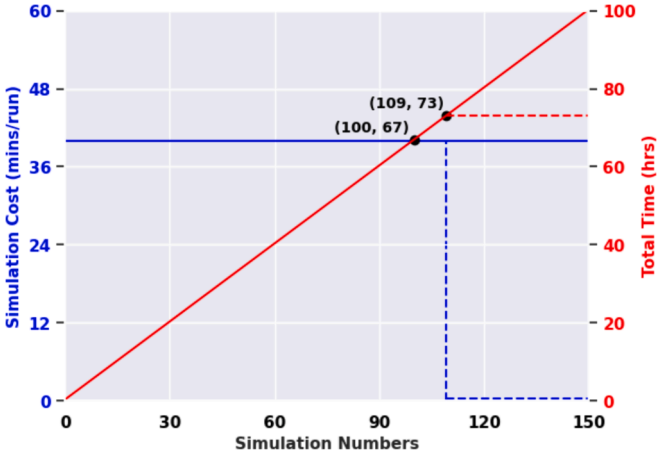


Fig. 10. Comparison between the computational costs of PFLOTRAN simulations and the cDC-GAN surrogate model. Running a numerical simulation on 12 CPU cores with PFLOTRAN takes about 40 min, as illustrated by the red solid line. Initially, for the first 100 simulations, both PFLOTRAN and cDC-GAN demand an equivalent amount of time. The cDC-GAN requires about 6 h on 4 GPU cores for training using the results of the preceding 100 PFLOTRAN simulations, equivalent to 9 model runs with PFLOTRAN. After the training of the cDC-GAN is completed, the computational cost becomes nearly minimal (represented by the red dashed line). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. Discussion

The results presented in this study demonstrate the effectiveness of the cDC-GAN model in predicting the spatial and temporal distributions of biogeochemical reaction rates and concentrations of reaction products in the river-aquifer system. Unlike traditional and computationally expensive RTMs, cDC-GAN can provide faster and more efficient solutions to the inherently nonlinear reactive flow and transport problem in the HZ. To ensure comprehensive evaluation, the models were individually trained for distinct heterogeneity scenarios characterized by different combinations of high-k and low-k facies with varying volume fractions and were then evaluated using a separate set of heterogeneity maps to assess their performance. The cDC-GAN models demonstrated their adaptability to handle various geological settings by capturing the spatial and temporal patterns of reaction rates and concentration plumes at different time steps. The qualitative analysis showed that in scenarios with a balanced distribution of high-k and low-k facies, the cDC-GAN exhibited optimal performance by effectively capturing plume details with minimal noise and artifacts, indicating higher fidelity in its predictions. Furthermore, the successful predictions of both the reaction zone size and the maximum reaction rates in all heterogeneity scenarios reaffirm the reliability of the cDC-GAN models in quantitative representation of the extent and intensity of biogeochemical reactions, including localized hotspots, within the river-aquifer system.

It is important to acknowledge certain limitations in this study and take them into account for future investigations. We focused on a simplified reaction scheme to establish a proof of concept and demonstrate the ability of the cDC-GAN models to simulate reactive solute transport within the hyporheic zone. Through this approach, we lay the groundwork for more comprehensive studies that consider more complex equilibrium and kinetically controlled reaction schemes and their interactions within the system, such as multiple concurrent reactions involving organic matter degradation, nutrient cycling, and redox reactions. Future studies could also explore the integration of additional environmental factors, such as temperature, pH, and other relevant variables to capture a more holistic understanding of the dynamic interplay between biogeochemical processes and solute transport within the hyporheic zone.

Moreover, this paper primarily focuses on the application of cDC-GAN in a 2D bimodal river-aquifer system to assess the feasibility of this proposed proxy model. One promising direction is to explore the application of cDC-GAN in multimodal hydrogeological systems, where multiple distinct flow regimes or components are present, such as interconnected aquifers with different hydraulic properties or the coexistence of rivers and fractured rock formations. Also, by extending the analysis to three-dimensional (3D) hydrogeological information such as varying lithology, aquifer properties, and boundary conditions, the application of cDC-GAN can be further evaluated for its effectiveness in simulating complex flow and transport processes in a more realistic setting. By addressing these considerations, we can advance our knowledge and develop more comprehensive surrogate models that capture the intricate dynamics of river-aquifer systems.

Another limitation of our work is the discrepancy between the

rectangular shape of the aquifer domains in our study, where the horizontal extent surpasses the vertical extent, and the typical training requirement of cDC-GANs for square-sized images. In architectures like U-net, the input images undergo multiple downscaling operations by a factor of two, which require power-of-two dimensions. As a solution, the input and output maps were resized to a uniform size of 128 × 128 before training the cDC-GAN models. However, this resizing process introduces the potential for changing map properties and the risk of losing crucial information through transformations. Such alterations may have a negative impact on the accuracy of the results obtained from the cDC-GAN models. One possible solution is to develop modified cDC-GAN architectures that can accommodate non-square input images while preserving the model's performance. Researchers can also investigate the use of data augmentation techniques tailored to handle rectangular-shaped aquifer domains and minimize the need for extensive resizing. Such advancements would enhance the accuracy and reliability of the cDC-GAN models for simulating reactive solute transport within non-square aquifer systems.

Due to their computational efficiency and ability to accurately predict reaction product concentrations and the distribution of reaction hotspots, cDC-GAN models like those developed in this work have significant potential for real-world applications. As the global population continues to rise and the associated demand on freshwater resources grows, increasing reliance on rivers and their adjacent aquifers necessitates a comprehensive understanding of water quality and patterns of contaminant transformations. Among the most significant contributions of cDC-GAN models is the potential to conduct data-worth analyses and explore which parameters are most influential on the hyporheic system. By informing practitioners and water resources managers of which hydrologic and geochemical properties their aquifer system is most sensitive too, cDC-GAN can save money and resources by better focusing field data collection efforts, and can thereby increase the efficiency with which water remediation strategies are implemented. Further, their predictive capacity beyond training datasets could make them invaluable tools for forecasting the efficacy of remedial actions. Future work should focus on the application of cDC-GAN models to real-world remediation planning to evaluate the amount of data necessary for their training and how their algorithms may be optimized for such purposes.

Additionally, it is essential to address the inherent limitation of interpretability common to many data-driven approaches. As a "black-box" model, the internal workings of the cDC-GAN model are opaque, hindering our understanding of how it learns and makes predictions from the input data. This lack of interpretability is a recognized challenge in the field of machine learning, particularly for complex models like GANs. Recent advances, such as physics-informed GANs (PI-GANs), have aimed to enhance interpretability by incorporating physical constraints and domain knowledge into the training process (Yang et al., 2019; Miele and Azevedo, 2024). By explicitly encoding known physical laws or principles into the loss function or architecture of the GAN, PI-GANs strive to generate outputs that are not only realistic but also consistent with the underlying physics. This approach has the potential to enhance the interpretability of the model by grounding its predictions in well-understood physical principles. However, the incorporation of physics-informed techniques into the cDC-GAN model introduce additional complexities and may require extensive domain expertise or supplementary data. Future research should explore strategies to improve the interpretability of the cDC-GAN model while retaining its computational efficiency, potentially leveraging insights from physics-informed approaches or other interpretability-enhancing techniques.

Finally, when comparing the GANs with other advanced generative deep learning models for subsurface modeling, it is important to consider their unique strengths and limitations. Compared to variational autoencoders (VAEs), the cDC-GAN model generates sharper and more realistic samples (Bao et al., 2022; Jozdani et al., 2022), which is essential for capturing the intricate details and heterogeneity of the subsurface system. Normalizing flow (NF) models can generate high-quality outputs but often require a large number of transformations to capture complex distributions, leading to increased computational complexity and memory requirements. (Bond-Taylor et al., 2021). Diffusion models have shown impressive results in generating high-quality samples but typically require a large number of iterative denoising steps during inference, which can be computationally expensive (Dhariwal and Nichol, 2021). In contrast, GANs can generate samples in a single forward pass once trained, making it more efficient for real-time applications or scenarios where rapid generation is required. GAN model stands out for its ability to generate realistic and detailed spatial patterns, its computational efficiency, and its flexibility in learning complex relationships from data. While other generative models have their own strengths and applications, the cDC-GAN model offers a balanced approach that combines the advantages of convolutional architectures, adversarial training, and conditional generation. As research in generative models continues to advance, exploring hybrid approaches that incorporate physical constraints and domain knowledge could further improve the accuracy and reliability of subsurface modeling predictions. By exploring a diverse range of modeling techniques, we can further advance our understanding and modeling capabilities in simulating biogeochemical processes in the hyporheic zone, contributing to more accurate and comprehensive assessments of environmental systems.

4. Conclusions

In this research, we employed a cDC-GAN surrogate model for simulating reactive transport in the hyporheic zone driven by river stage fluctuations. Our main objective was to accurately predict the spatial and temporal distribution of biogeochemical reaction rates and concentrations in heterogeneous bimodal aquifers. Traditionally, addressing the underlying flow and transport problem requires computationally intensive numerical simulations due to its highly nonlinear nature. However, our results demonstrate that cDC-GAN can effectively learn the complex nonlinear relationships between the input and output domains. Even when the heterogeneity structure of the aquifer changes, such as variations in the fractions of high-k and low-k facies, the cDC-GAN model maintains its robust performance. One notable advantage of the cDC-GAN model is its flexibility in modeling systems without making strong assumptions about linearity or relying on specific probability distribution functions. Instead, it primarily relies on the available data or patterns to learn and understand the underlying relationships between the input and output variables. This attribute empowers the cDC-GAN model to adapt and capture the dynamics of diverse geological settings, thus proving its applicability across a broad range of scenarios.

5. Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to improve language and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

CRediT authorship contribution statement

Farzad Moeini: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Reza Ershadnia:** Writing – review & editing, Methodology. **Rebecca L. Rubinstein:** Writing – review & editing, Software, Methodology. **Roelof Versteeg:** Writing – review & editing, Software, Methodology. **Pei Li:** Writing – review & editing. **Jeffery T. McGarr:** Writing – review & editing. **Alireza Meyal:** Software, Methodology. **Corey D. Wallace:** Writing – review & editing. **Zhenxue Dai:** Writing – review & editing. **Kenneth C. Carroll:** Writing

– review & editing. **Mohamad Reza Soltanian:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

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

Data availability

We have shared the link to part of our data. Full data will be available upon request.

Acknowledgement

This work was supported by the U.S. Department of Energy under SBIR award DE-SC0019621 to Subsurface Insights LLC. The fifth author received support from the National Science Foundation under grant EAR-2048452. Any opinions, findings, conclusions, or recommendations expressed in this article are solely those of the authors and do not necessarily reflect the views of the National Science Fosundation.The authors gratefully acknowledge the Ohio Supercomputer Center (OSC) for providing the computational resources necessary for this study. The authors would also like to extend their gratitude to Mr. Aladdin Persson for providing the foundational DC-GAN Python code, which was modified for this research. The original code can be found at: https://github.com/aladdinpersson/Machine-Learning-Collection/tree/master/ML/Pytorch/GANs/2. DCGAN. The training data and PFLOTRAN input file and are provided at https://doi.org/10.5281/zenodo.11090450.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2024.131485.

References

Aggarwal, A., Mittal, M., Battineni, G., 2021. Generative adversarial network: An overview of theory and applications. Int. J. Info. Manage. Data Insights 1 (1), 100004. https://doi.org/10.1016/j.jjimei.2020.100004.

Akhtar, Naseem, Muhammad I. Syakir, Mardiana Idayu Ahmad, Mohd Talha Anees, Ahmad Farid Bin Abu Bakar, Syed Adil Mizan, Sami Farraj Alsaadi, Mohammad Muqtada Ali Khan, and Mohamad Shaiful Md Yusuff. "Upscaling of surface water and groundwater interactions in hyporheic zone from local to regional scale."*Water* 14, no. 4 (2022): 647. DOI: https://doi.org/10.3390/w14040647.

Alqahtani, H., Kavakli-Thorne, M., Kumar, G., 2021. Applications of generative adversarial networks (gans): An updated review. Arch. Comput. Meth. Eng. 28, 525–552. https://doi.org/10.1007/s11831-019-09388-y.

Anderson, D.M., Glibert, P.M., Burkholder, J.M., 2002. Harmful algal blooms and eutrophication: nutrient sources, composition, and consequences. Estuaries 25, 704–726. https://doi.org/10.1007/BF02804901.

Asher, Michael J., Barry FW Croke, Anthony J. Jakeman, and Luk JM Peeters. "A review of surrogate models and their application to groundwater modeling." *Water Resources Research* 51, no. 8 (2015): 5957-5973. DOI: https://doi.org/10.1002/2015WR016967.

Bao, J., Li, L., Davis, A., 2022. Variational autoencoder or generative adversarial networks? A comparison of two deep learning methods for flow and transport data assimilation. Math. Geosci. 54 (6), 1017–1042. https://doi.org/10.1007/s11004-022-10003-3.

Bause, M., Knabner, P., 2004. Computation of variably saturated subsurface flow by adaptive mixed hybrid finite element methods. Adv. Water Resour. 27 (6), 565–581. https://doi.org/10.1016/j.advwatres.2004.03.005.

Bear, J., 1972. Dynamics of fluids in porous media. Elsevier http://refhub.elsevier.com/S0309-1708(21)00180-9/sbref0005.

Bed form-induced hyporheic exchange and geochemical hotspots F. Behzadi C.D. Wallace D. Ward H. Zhou R. Versteeg M.R. Soltanian Advances in Water Resources 156 2021 104025 10.1016/j.advwatres.2021.104025.

Bell, P.R.F., 1992. Eutrophication and coral reefs—some examples in the Great Barrier Reef lagoon. Water Res. 26 (5), 553–568. https://doi.org/10.1016/0043-1354(92)90228-V.

Bernhardt, E.S., Blaszczak, J.R., Ficken, C.D., Fork, M.L., Kaiser, K.E., Seybold, E.C., 2017. Control points in ecosystems: moving beyond the hot spot hot moment concept. Ecosystems 20, 665–682. https://doi.org/10.1007/s10021-016-0103-y.

Birch, G.F., Taylor, S.E., Matthai, C., 2001. Small-scale spatial and temporal variance in the concentration of heavy metals in aquatic sediments: a review and some new concepts. Environ. Pollut. 113 (3), 357–372. https://doi.org/10.1016/S0269-7491(00)00182-2.

Boano, F., Demaria, A., Revelli, R., Ridolfi, L., 2010. "Biogeochemical zonation due to intrameander hyporheic flow.". Water Res. Res. 46 (2) https://doi.org/10.1029/2008WR007583.

Boano, F., Harvey, J.W., Marion, A., Packman, A.I., Revelli, R., Ridolfi, L., Wörman, A., 2014. Hyporheic flow and transport processes: Mechanisms, models, and biogeochemical implications. Rev. Geophys. 52 (4), 603–679. https://doi.org/10.1002/2012RG000417.

Bond-Taylor, S., Leach, A., Long, Y., Willcocks, C.G., 2021. Deep generative modelling: A comparative review of VAEs, GANs, normalizing flows, energy-based and autoregressive models. IEEE Trans. Pattern Anal. Mach. Intell. 44 (11), 7327–7347. https://doi.org/10.1109/TPAMI.2021.3116668.

Boulton, A.J., Findlay, S., Marmonier, P., Stanley, E.H., Maurice Valett, H., 1998. The functional significance of the hyporheic zone in streams and rivers. In: Annual Review of Ecology and Systematics 29(1), pp. 59–81. https://doi.org/10.1146/annurev.ecolsys.29.1.59.

Brophy, E., Wang, Z., She, Q.i., Ward, T., 2023. Generative adversarial networks in time series: A systematic literature review. ACM Comput. Surv. 55 (10), 1–31. https://doi.org/10.1145/3559540.

Brunner, Philip, René Therrien, Philippe Renard, Craig T. Simmons, and Harrie-Jan Hendricks Franssen. "Advances in understanding river-groundwater interactions." *Reviews of Geophysics* 55, no. 3 (2017): 818-854. DOI: https://doi.org/10.1002/2017RG000556.

Cardenas, M.B., Wilson, J.L., Zlotnik, V.A., 2004. Impact of heterogeneity, bed forms, and stream curvature on subchannel hyporheic exchange. Water Resour. Res. 40, no. 8 https://doi.org/10.1029/2004WR003008.

Cardenas, M. Bayani. "Hyporheic zone hydrologic science: A historical account of its emergence and a prospectus." *Water Resources Research* 51, no. 5 (2015): 3601-3616. DOI: https://doi.org/10.1002/2015WR017028.

Carle, S.F., 1999. T-PROGS: transition probability geostatistical software, version 2.1. Department of Land, Air and Water Resources. University of California, Davis.

Chan, Shing, and Ahmed H. Elsheikh. "Parametrization and generation of geological models with generative adversarial networks." arXiv preprint arXiv:1708.01810 (2017). DOI:.

Chang, F.-J., Lin, C.-H., Chang, K.-C., Kao, Y.-H., Chang, L.-C., 2014. Investigating the interactive mechanisms between surface water and groundwater over the Jhuoshuei river basin in central Taiwan. Paddy Water Environ, 12, 365–377. https://doi.org/10.1007/s10333-013-0391-1.

Claret, C., Boulton, A.J., 2009. Integrating hydraulic conductivity with biogeochemical gradients and microbial activity along river-groundwater exchange zones in a subtropical stream. Hydrgeol. J. 17 (1), 151. https://doi.org/10.1007/s10040-008-0373-3.

Clément, J.-B., Sous, D., Bouchette, F., Golay, F., Ersoy, M., 2023. A Richards' equation-based model for wave-resolving simulation of variably-saturated beach groundwater flow dynamics. J. Hydrol. 619, 129344 https://doi.org/10.1016/j.jhydrol.2023.129344.

Clement, T.P., Wise, W.R., Molz, F.J., 1994. A physically based, two-dimensional, finite-difference algorithm for modeling variably saturated flow. Journal of Hydrology 161 (1–4) 71–90. https://doi.org/10.1016/0022-1694(94)90121-X.

Dhariwal, P., Nichol, A., 2021. Diffusion models beat GANs on image synthesis. Adv. Neural Inf. Proces. Syst. 34, 8780–8794. https://doi.org/10.48550/arXiv.2105.05233.

Dwivedi, D., Arora, B., Steefel, C.I., Dafflon, B., Versteeg, R., 2018. Hot spots and hot moments of nitrogen in a riparian corridor. Water Resour. Res. 54 (1), 205–222. https://doi.org/10.1002/2017WR022346.

Dwivedi, Dipankar, Carl I. Steefel, Bhavna Arora, Michelle Newcomer, J. David Moulton, Baptiste Dafflon, Boris Faybishenko et al. "Geochemical exports to river from the intrameander hyporheic zone under transient hydrologic conditions: East River Mountainous Watershed, Colorado." *Water Resources Research* 54, no. 10 (2018a): 8456-8477. DOI: https://doi.org/10.1029/2018WR023377.

Elshall, A.S., Arik, A.D., El-Kadi, A.I., Pierce, S., Ye, M., Burnett, K.M., Wada, C.A., Bremer, L.L., Chun, G., 2020. Groundwater sustainability: A review of the interactions between science and policy. Environ. Res. Lett. 15 (9), 093004 https://doi.org/10.1088/1748-9326/ab8e8c.

Frei, S., Fleckenstein, J.H., Kollet, S.J., Maxwell, R.M., 2009. Patterns and dynamics of river–aquifer exchange with variably-saturated flow using a fully-coupled model. J. Hydrol. 375 (3–4), 383–393. https://doi.org/10.1016/j.jhydrol.2009.06.038.

Gillespie, D.T., 2000. The chemical Langevin equation. J. Chem. Phys. 113 (1), 297–306. https://doi.org/10.1063/1.481811.

Gonog, L., Zhou, Y., 2019. A review: generative adversarial networks. In: In *2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, pp. 505–510. https://doi.org/10.1109/ICIEA.2019.8833686.

Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative adversarial nets." *Advances in neural information processing systems* 27 (2014). DOI: https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf.

Gramling, C.M., Harvey, C.F., Meigs, L.C., 2002. Reactive transport in porous media: A comparison of model prediction with laboratory visualization. Environ. Sci. Tech. 36 (11), 2508–2514. https://doi.org/10.1021/es0157144.

Gu, C., Anderson, W., Maggi, F., 2012. Riparian biogeochemical hot moments induced by stream fluctuations. Water Resour. Res. 48, no. 9 https://doi.org/10.1029/2011WR011720.

Harvey, J.W., Böhlke, J.K., Voytek, M.A., Scott, D., Tobias, C.R., 2013. Hyporheic zone denitrification: Controls on effective reaction depth and contribution to whole-stream mass balance. Water Resour. Res. 49 (10), 6298–6316. https://doi.org/10.1002/wrcr.20492.

Hester, E.T., Bayani Cardenas, M., Haggerty, R., Apte, S.V., 2017. The importance and challenge of hyporheic mixing. Water Resour. Res. 53 (5), 3565–3575. https://doi.org/10.1002/2016WR020005.

Hester, E.T., Santizo, K.Y., Nida, A.A., Widdowson, M.A., 2021. Hyporheic transverse mixing zones and dispersivity: Laboratory and numerical experiments of hydraulic controls. J. Contam. Hydrol. 243, 103885 https://doi.org/10.1016/j.jconhyd.2021.103885.

Hunt, R.J., Doherty, J., Tonkin, M.J., 2007. Are models too simple? Arguments for increased parameterization. Groundwater 45 (3), 254–262. https://doi.org/10.1111/j.1745-6584.2007.00316.x.

Irvine, D.J., Cranswick, R.H., Simmons, C.T., Shanafield, M.A., Lautz, L.K., 2015. The effect of streambed heterogeneity on groundwater-surface water exchange fluxes inferred from temperature time series. Water Resour. Res. 51 (1), 198–212. https://doi.org/10.1002/2014WR015769.

Isola, Phillip, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. "Image-to-image translation with conditional adversarial networks." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125-1134. 2017. DOI: https://doi.org/10.48550/arXiv.1611.07004.

Jan, A., Coon, E.T., Painter, S.L., 2021. Toward more mechanistic representations of biogeochemical processes in river networks: Implementation and demonstration of a multiscale model. Environ. Model. Softw. 145, 105166 https://doi.org/10.1016/j.envsoft.2021.105166.

Jazayeri, A., Werner, A.D., Cartwright, N., 2021. Partially penetrating lake-aquifer interaction in a laboratory-scale tidal setting. J. Hydrol. 603, 127080 https://doi.org/10.1016/j.jhydrol.2021.127080.

Jordan, M.I., Mitchell, T.M., 2015. Machine learning: Trends, perspectives, and prospects. Science 349 (6245), 255–260. https://doi.org/10.1126/science.aaa8415.

Jozdani, S., Chen, D., Pouliot, D., Johnson, B.A., 2022. A review and meta-analysis of generative adversarial networks and their applications in remote sensing. Int. J. Appl. Earth Obs. Geoinf. 108, 102734 https://doi.org/10.1016/j.jag.2022.102734.

Kirkland, M.R., Hills, R.G., Wierenga, P.J., 1992. Algorithms for solving Richards' equation for variably saturated soils. Water Resour. Res. 28 (8), 2049–2058. https://doi.org/10.1029/92WR00802.

Krause, S., Abbott, B.W., Baranov, V., Bernal, S., Blaen, P., Datry, T., Drummond, J., et al., 2022. Organizational principles of hyporheic exchange flow and biogeochemical cycling in river networks across scales. Water Resour. Res. 58 (3), e2021WR029771 https://doi.org/10.1029/2021WR029771.

Lautz, L.K., Fanelli, R.M., 2008. Seasonal biogeochemical hotspots in the streambed around restoration structures. Biogeochemistry 91, 85–104. https://doi.org/10.1007/s10533-008-9235-2.

Lautz, L.K., Siegel, D.I., 2006. Modeling surface and ground water mixing in the hyporheic zone using MODFLOW and MT3D. Adv. Water Resour. 29 (11), 1618–1633. https://doi.org/10.1016/j.advwatres.2005.12.003.

Ledig, Christian, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken et al. "Photo-realistic single image super-resolution using a generative adversarial network." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4681-4690. 2017. DOI: https://doi.org/10.48550/arXiv.1609.04802.

Lewandowski, J., Meinikmann, K., Krause, S., 2020. Groundwater–surface water interactions: Recent advances and interdisciplinary challenges. Water 12 (1), 296. https://doi.org/10.3390/w12010296.

Li, P., Berkowitz, B., 2019. Characterization of mixing and reaction between chemical species during cycles of drainage and imbibition in porous media. Adv. Water Resour. 130, 113–128. https://doi.org/10.1016/j.advwatres.2019.06.003.

Li, B., Liu, X., Kaufman, M.H., Turetcaia, A., Chen, X., Bayani Cardenas, M., 2020. Flexible and modular simultaneous modeling of flow and reactive transport in rivers and hyporheic zones. Water Resour. Res. 56, no. 2, e2019WR026528 https://doi.org/10.1029/2019WR026528.

Li, L.i., Maher, K., Navarre-Sitchler, A., Druhan, J., Meile, C., Lawrence, C., Moore, J., et al., 2017. Expanding the role of reactive transport models in critical zone processes. Earth Sci. Rev. 165, 280–301. https://doi.org/10.1016/j.earscirev.2016.09.001.

McClain, M.E., Boyer, E.W., Lisa Dent, C., Gergel, S.E., Grimm, N.B., Groffman, P.M., Hart, S.C., et al., 2003. Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. Ecosystems 301–312. https://doi.org/10.1007/s10021-003-0161-9.

Miele, R., Azevedo, L., 2024. Physics-informed W-Net GAN for the direct stochastic inversion of fullstack seismic data into facies models. Sci. Rep. 14 (1), 5122. https://doi.org/10.1038/s41598-024-55683-5.

Mirza, Mehdi, and Simon Osindero. "Conditional generative adversarial nets." *arXiv preprint arXiv:1411.1784* (2014). DOI: https://doi.org/10.48550/arXiv.1411.1784.

Mori, N., Debeljak, B., Škerjanec, M., Simčič, T., Kanduč, T., Brancelj, A., 2019. Modelling the effects of multiple stressors on respiration and microbial biomass in the hyporheic zone using decision trees. Water Res. 149, 9–20. https://doi.org/10.1016/j.watres.2018.10.093.

Mosser, L., Dubrule, O., Blunt, M.J., 2017. Reconstruction of three-dimensional porous media using generative adversarial neural networks. Phys. Rev. E 96 (4), 043309. https://doi.org/10.1103/PhysRevE.96.043309.

Ng, G.-H., Rosenfeld, C.E., Santelli, C.M., Yourd, A.R., Lange, J., Duhn, K., Johnson, N. W., 2020. Microbial and reactive transport modeling evidence for hyporheic flux-driven cryptic sulfur cycling and anaerobic methane oxidation in a sulfate-impacted wetland-stream system. J. Geophys. Res. Biogeo. 125, no. 2, e2019JG005185 https://doi.org/10.1029/2019JG005185.

Painter, S.L., 2021. On the representation of hyporheic exchange in models for reactive transport in stream and river corridors. Frontiers in Water 2, 595538. https://doi.org/10.3389/frwa.2020.595538.

Palumbo-Roe, B., Wragg, J., Banks, V.J., 2012. Lead mobilisation in the hyporheic zone and river bank sediments of a contaminated stream: contribution to diffuse pollution. J. Soil. Sediment. 12, 1633–1640. https://doi.org/10.1007/s11368-012-0552-7.

Perzan, Z., Babey, T., Caers, J., Bargar, J.R., Maher, K., 2021. Local and global sensitivity analysis of a reactive transport model simulating floodplain redox cycling. Water Resour. Res. 57 (12), e2021WR029723 https://doi.org/10.1029/2021WR029723.

Pryshlak, T.T., Sawyer, A.H., Stonedahl, S.H., Soltanian, M.R., 2015. Multiscale hyporheic exchange through strongly heterogeneous sediments. Water Resour. Res. 51 (11), 9127–9140. https://doi.org/10.1002/2015WR017293.

Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." *arXiv preprint arXiv: 1511.06434* (2015). DOI: https://doi.org/10.48550/arXiv.1511.06434.

Razavi, S., Tolson, B.A., Burn, D.H., 2012. Review of surrogate modeling in water resources. Water Resour. Res. 48, no. 7 https://doi.org/10.1029/2011WR011527.

Ren, H., Song, X., Yilin Fang, Z., Hou, J., Scheibe, T.D., 2021. Machine learning analysis of hydrologic exchange flows and transit time distributions in a large regulated river. Frontiers in Artificial Intelligence 4, 648071. https://doi.org/10.3389/frai.2021.648071.

Richards, L.A., 1931. Capillary conduction of liquids through porous mediums. Physics 1 (5), 318–333. https://doi.org/10.1063/1.1745010.

Russel, S., 2010. and P. Pearson Education Inc, Norvig. Artificial intelligence a modern approach.

Ryan, R.J., Boufadel, M.C., 2006. Influence of streambed hydraulic conductivity on solute exchange with the hyporheic zone. Environ. Geol. 51, 203–210. https://doi.org/10.1007/s00254-006-0319-9.

Salehin, M., Packman, A.I., Paradis, M., 2004. Hyporheic exchange with heterogeneous streambeds: Laboratory experiments and modeling. Water Resour. Res. 40, no. 11 https://doi.org/10.1029/2003WR002567.

Sawyer, A.H., Bayani Cardenas, M., 2009. Hyporheic flow and residence time distributions in heterogeneous cross-bedded sediment. Water Resour. Res. 45, no. 8 https://doi.org/10.1029/2008WR007632.

Scanlon, B.R., Fakhreddine, S., Rateb, A., de Graaf, I., Famiglietti, J., Tom Gleeson, R., Grafton, Q., et al., 2023. Global water resources and the role of groundwater in a resilient water future. Nature Reviews Earth & Environment 4 (2), 87–101. https://doi.org/10.1038/s43017-022-00378-6.

Shen, C., Chen, X., Laloy, E., 2021. "Editorial: broadening the use of machine learning in hydrology. Frontiers in Water 3. https://doi.org/10.3389/frwa.2021.681023.

Singh, T., Gupta, S., Chiogna, G., Krause, S., Wohlmuth, B., 2022. Impacts of peak-flow events on hyporheic denitrification potential. Water Resources Research 58 (3), e2021WR031407. https://doi.org/10.1029/2021WR031407.

Soltanian, Mohamad Reza, Robert W. Ritzi, Chao Cheng Huang, and Zhenxue Dai. "Relating reactive solute transport to hierarchical and multiscale sedimentary architecture in a Lagrangian-based transport model: 1. Time-dependent effective retardation factor." *Water Resources Research* 51, no. 3 (2015a): 1586-1600. DOI: https://doi.org/10.1002/2014WR016353.

Soltanian, Mohamad Reza, Robert W. Ritzi, Chao Cheng Huang, and Zhenxue Dai. "Relating reactive solute transport to hierarchical and multiscale sedimentary architecture in a L agrangian-based transport model: 2. Particle displacement variance." *Water Resources Research* 51, no. 3 (2015b): 1601-1618. DOI: https://doi.org/10.1002/2014WR016354.

Song, S., Mukerji, T., Hou, J., 2021. Bridging the gap between geophysics and geology with generative adversarial networks. IEEE Trans. Geosci. Remote Sens. 60, 1–11. https://doi.org/10.1109/TGRS.2021.3066975.

Stegen, J.C., Fredrickson, J.K., Wilkins, M.J., Konopka, A.E., Nelson, W.C., Arntzen, E.V., Chrisler, W.B., et al., 2016. Groundwater–surface water mixing shifts ecological assembly processes and stimulates organic carbon turnover. Nat. Commun. 7 (1), 11237. https://doi.org/10.1038/ncomms11237.

Stegen, J.C., Johnson, T., Fredrickson, J.K., Wilkins, M.J., Konopka, A.E., Nelson, W.C., Arntzen, E.V., et al., 2018. Influences of organic carbon speciation on hyporheic corridor biogeochemistry and microbial ecology. Nat. Commun. 9 (1), 585. https://doi.org/10.1038/s41467-018-02922-9.

Suk, H., Park, E., 2019. Numerical solution of the Kirchhoff-transformed Richards equation for simulating variably saturated flow in heterogeneous layered porous media. J. Hydrol. 579, 124213 https://doi.org/10.1016/j.jhydrol.2019.124213.

Tonina, D., de Barros, F.PJ., Marzadri, A., Bellin, A., 2016. Does streambed heterogeneity matter for hyporheic residence time distribution in sand-bedded streams? Adv. Water Resour. 96, 120–126. https://doi.org/10.1016/j.advwatres.2016.07.009.

Trauth, N., Schmidt, C., Vieweg, M., Maier, U., Fleckenstein, J.H., 2014. Hyporheic transport and biogeochemical reactions in pool-riffle systems under varying ambient groundwater flow conditions. J. Geophys. Res. Biogeo. 119 (5), 910–928. https://doi.org/10.1002/2013JG002586.

Van Genuchten, M.T., 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.

Wallace, C.D., Sawyer, A.H., Soltanian, M.R., Barnes, R.T., 2020. Nitrate removal within heterogeneous riparian aquifers under tidal influence. Geophys. Res. Lett. 47 (10), e2019GL085699 https://doi.org/10.1029/2019GL085699.

Wallis, I., Prommer, H., Berg, M., Siade, A.J., Sun, J., Kipfer, R., 2020. The river–groundwater interface as a hotspot for arsenic release. Nat. Geosci. 13 (4), 288–295. https://doi.org/10.1038/s41561-020-0557-6.

Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 13 (4), 600–612. https://doi.org/10.1109/TIP.2003.819861.

Western, A.W., Blöschl, G., Grayson, R.B., 2001. Toward capturing hydrologically significant connectivity in spatial patterns. Water Resour. Res. 37 (1), 83–97. https://doi.org/10.1029/2000WR900241.

Wu, X., Ma, T., Wang, Y., 2020. Surface water and groundwater interactions in wetlands. J. Earth Sci. 31, 1016–1028. https://doi.org/10.1007/s12583-020-1333-7.

Yabusaki, S.B., Wilkins, M.J., Fang, Y., Williams, K.H., Arora, B., Bargar, J., Beller, H.R., et al., 2017. Water table dynamics and biogeochemical cycling in a shallow, variably-saturated floodplain. Environ. Sci. Tech. 51 (6), 3307–3317. https://doi.org/10.1021/acs.est.6b04873.

Yang, L., Treichler, S., Kurth, T., Fischer, K., Barajas-Solano, D., Romero, J., Churavy, V., et al., 2019b. Highly-scalable, physics-informed GANs for learning solutions of stochastic PDEs. In: In 2019 IEEE/ACM Third Workshop on Deep Learning on Supercomputers (DLS). IEEE, pp. 1–11. https://doi.org/10.1109/DLS49591.2019.00006.

Yang, C., Zheng, F., Liu, Y., Zhang, Y.-K., Liu, W., Zhang, Q., Yang, X., 2019a. Modeling hydro-biogeochemical transformation of chromium in hyporheic zone: Effects of spatial and temporal resolutions. J. Hydrol. 579, 124152 https://doi.org/10.1016/j.jhydrol.2019.124152.

Zhan, C., Dai, Z., Samper, J., Yin, S., Ershadnia, R., Zhang, X., Wang, Y., Yang, Z., Luan, X., Soltanian, M.R., 2022. An integrated inversion framework for heterogeneous aquifer structure identification with single-sample generative adversarial network. J. Hydrol. 610, 127844 https://doi.org/10.1016/j.jhydrol.2022.127844.

Zhang, Kai, Luc Van Gool, and Radu Timofte. "Deep unfolding network for image super-resolution." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 3217-3226. 2020. DOI: https://doi.org/10.48550/arXiv.2003.10428.

Zhang, Kaifeng. "On mode collapse in generative adversarial networks." In *Artificial Neural Networks and Machine Learning–ICANN 2021: 30th International Conference on Artificial Neural Networks, Bratislava, Slovakia, September 14–17, 2021, Proceedings, Part II 30*, pp. 563-574. Springer International Publishing, 2021. DOI: https://doi.org/10.1007/978-3-030-86340-1_45.

Zhong, Z., Sun, A.Y., Jeong, H., 2019. Predicting CO_2 plume migration in heterogeneous formations using conditional deep convolutional generative adversarial network. Water Resour. Res. 55 (7), 5830–5851. https://doi.org/10.1029/2018WR024592.

Zhong, Z., Sun, A.Y., Wang, Y., Ren, B.o., 2020. Predicting field production rates for waterflooding using a machine learning-based proxy model. J. Pet. Sci. Eng. 194, 107574 https://doi.org/10.1016/j.petrol.2020.107574.