

Physics-Informed Neural Networks for Seepage Modeling in Porous Media: A Review

Baihua Zeng

*School of Civil Engineering, Tianjin University, Tianjin, China
barryzeng0806@gmail.com*

Abstract. Seepage in porous media governs groundwater development, the safety of dams and levees, oil, gas, and geothermal extraction, geological CO₂ storage, and water migration in the vadose zone. Finite difference, finite element, finite volume, and lattice Boltzmann methods remain the mainstream tools for engineering forward modeling because they are supported by mature discretization theory, conservation control, and verifiable numerical stability. However, when boundary conditions are incomplete, medium parameters are difficult to measure, constitutive relations are uncertain, or observations are sparse, traditional workflows often require laborious mesh reconstruction, repeated forward simulations, and strong prior regularization. Physics-informed neural networks (PINNs) provide a unified differentiable modeling framework for seepage forward modeling, inversion, and state reconstruction by embedding governing equations, initial and boundary conditions, observational data, and structural priors simultaneously in the training loss. This review surveys the field from two perspectives. First, it traces the methodological evolution of PINNs from joint parameter-state identification to monotonicity constraints, domain decomposition, weak-form/conservative residuals, mixed pressure-velocity formulations, explicit fracture embedding, and neural operators. Second, it distinguishes forward problem solution, inverse modeling, and flow-field reconstruction by task type. Existing studies show that PINNs are most advantageous for sparse-data inversion and physics-constrained state completion. In large-scale, strongly nonlinear, three-dimensional engineering forward modeling, they are better regarded as differentiable complements to conventional solvers rather than as universal replacements. Future progress will depend on conservative discrete residuals, hybrid PINN-numerical-solver frameworks, adaptive sampling, identifiability and uncertainty quantification, and digital-twin workflows built around monitoring data.

Keywords: physics-informed neural networks, porous media, seepage, inverse modeling, uncertainty quantification

1. Introduction

Seepage in porous media is a fundamental process in groundwater hydrology, geotechnical engineering, reservoir engineering, oil-reservoir development, contaminant transport, and geological CO₂ storage. Its state evolution is jointly controlled by pore structure, heterogeneity, saturation, capillary pressure, fracture networks, and coupled thermal-hydraulic-mechanical-chemical

processes. At the continuum scale, Darcy's equation, Richards' equation, and multiphase conservation equations are commonly used models. At the pore scale, the Stokes equations or lattice Boltzmann methods are used to resolve local flow in complex pore geometries [1-3]. Conventional numerical methods retain irreplaceable advantages in these problems: the finite element method (FEM) is well suited to complex boundaries and variational formulations; the finite volume method (FVM) is well suited to local conservation and multiphase flow; the finite difference method (FDM) is efficient on regular grids and in groundwater models; and the lattice Boltzmann method (LBM) is suitable for complex pore-scale boundaries.

What has truly brought physics-informed neural networks (PINNs) into seepage research is not the failure of conventional solvers for standard forward problems, but the long-standing lack of an efficient, unified, differentiable framework for seepage inverse problems and monitoring problems. Engineering sites usually provide only limited data on hydraulic head, pressure, saturation, water content, or electrical resistivity, whereas hydraulic conductivity, relative permeability, capillary-pressure curves, well source/sink terms, and boundary fluxes determine the predictive capability of the model. Traditional inversion generally requires repeated calls to forward models and relies on prior parameterization and regularization. When the unknowns include both state fields and constitutive relations, non-uniqueness becomes even more pronounced. The core value of PINNs is that they place partial differential equation (PDE) residuals, observation errors, initial and boundary conditions, and physical priors in the same optimization problem, allowing state estimation, parameter identification, and constitutive learning to proceed in parallel [4-7]. The main flow processes, methodological advances, and engineering applications reviewed in this paper are summarized in Figure 1.

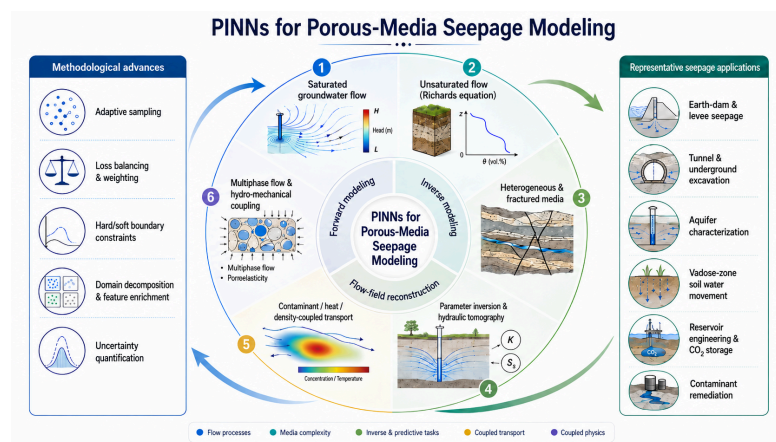


Figure 1. Overview of PINN-based seepage modeling in porous media

Accordingly, this review does not describe PINNs as "mesh-free replacement solvers"; rather, it treats them as differentiable data-physics fusion tools. Unlike most reviews organized by application area, this paper follows two threads. The first is a historical thread, explaining why network architectures, residual forms, and training strategies have changed with the specific difficulties of seepage problems. The second is a task-oriented thread, classifying the literature into forward problem solution, inverse modeling, and flow-field reconstruction according to the nature of the unknowns. This organization avoids simply juxtaposing numerous case studies and more clearly addresses an engineering question: under what conditions are PINNs valuable, and under what conditions should established numerical models still be relied upon?

2. Review framework: physics, unknowns and tasks

The loss function of a seepage PINN is typically composed of four types of terms: governing-equation residuals, initial/boundary-condition residuals, observational-data errors, and structural priors. For saturated flow, the core constraints arise from mass conservation and Darcy's law. For unsaturated flow, Richards' equation further requires the water retention curve and the unsaturated hydraulic conductivity function [8-11]. For multiphase flow, phase conservation, capillary pressure, relative permeability, well source/sink terms, and interfacial fronts must also be handled. In a classical strong-form PINN, the network directly approximates pressure, hydraulic head, or saturation, and PDE residuals are computed by automatic differentiation [12]. In weak-form, mixed-variable, or finite-volume residual formulations, fluxes, conservation integrals, or discrete cell-to-cell exchanges are explicitly written into the training objective [13-16].

From the task perspective, seepage PINNs can be divided into three categories. Forward problem solution assumes that the equations, parameters, and boundaries are known, and aims to predict complete pressure, hydraulic-head, saturation, or flux fields. Inverse modeling takes partial observations as input and seeks to identify hydraulic conductivity, constitutive relations, source/sink terms, boundary fluxes, or fracture parameters. Flow-field reconstruction lies between the two: it does not necessarily pursue complete parameter inversion, but reconstructs the spatiotemporal state of unobserved regions under sparse observations and physical constraints. The three task types often overlap, but their dominant unknowns and evaluation criteria differ. Forward problems emphasize error, conservation, and scalability. Inverse problems emphasize identifiability, physical plausibility of priors, and independent validation. Reconstruction tasks emphasize observational coverage, physical consistency, and uncertainty coverage. Table 1 summarizes the corresponding task taxonomy.

Table 1. Task taxonomy of PINNs in porous media seepage

Task	Primary unknowns	Typical inputs	Key evaluation criteria
Forward problem solution	Pressure, hydraulic head, saturation, flux	PDEs, parameters, boundary and initial conditions	Error, local conservation, computational efficiency
Inverse modeling	Hydraulic conductivity, constitutive relations, source/sink terms, boundary fluxes	Sparse observations + PDEs + prior information	Identifiability, robustness, independent validation
Flow-field reconstruction	State fields in unobserved regions	Sparse/noisy observations + PDEs	Reconstruction error, residuals, consistency, uncertainty

3. Methodological evolution of PINNs in seepage modeling

The first stage focused on inverse problems. The PINN paradigm proposed by Raissi et al. unified PDEs, initial and boundary conditions, and observation errors into a differentiable loss. Tartakovsky et al. then applied it to parameter and constitutive-relation learning in subsurface flow: for saturated flow, neural networks were used to represent hydraulic conductivity and hydraulic head separately; for unsaturated flow, the conductivity function and capillary-pressure relation were further learned [4, 8]. The importance of this work lies in its expansion of the unknowns from a single state variable to parameter fields and constitutive functions, thereby changing the traditional serial workflow of "specifying a constitutive law - forward modeling - calibration." He et al.'s multiphysics PINN

(MPINN) further incorporated hydraulic head, concentration, and hydraulic conductivity into a multiphysics data assimilation framework, showing that multi-source observations can improve joint parameter-state identification [9].

The limitations of this stage were also clear: most examples were low-dimensional, geometrically simple, and involved relatively smooth media. Although PDE residuals provided physical constraints, they did not automatically guarantee credible constitutive forms, interfacial fluxes, or local conservation. Subsequent studies therefore turned to structural priors and reliability-oriented modifications.

The second stage shifted toward reliability improvements, with particular emphasis on non-smoothness and ill-posedness. Constitutive curves in unsaturated flow have clear monotonicity and boundary behavior. If they are regressed using an ordinary multilayer perceptron (MLP) alone, the resulting function may satisfy the PDE mathematically while remaining physically unreasonable. The monotonicity-constrained Richardson-Richards equation PINN (RRE-PINN) proposed by Bandai and Ghezzehei encoded the physical shapes of the water retention curve and the hydraulic conductivity function into the network architecture, allowing volumetric-water-content observations to constrain the state, constitutive relations, and water flux simultaneously [10]. Its significance lies in moving from merely "satisfying the governing equation" to satisfying the governing equation while also respecting constitutive priors.

At the same time, discontinuities in the medium motivated domain decomposition. For layered soils or hydraulic-conductivity jumps, a single smooth network covering the entire domain introduces an erroneous global smoothing bias. Domain-decomposition PINNs assign separate networks to different subdomains and couple them through interface conditions, which better matches the structure of piecewise smooth solutions [11]. The groundwater PINN (GW-PINN) further showed that high gradients near wells or source/sink terms cannot be handled adequately by uniform collocation; locally refined sampling and time-decomposition training can substantially reduce training difficulty [13]. The central change in this stage was that PINN reliability no longer depended on "more training epochs," but on correct physical priors, subdomain organization, and the allocation of sampling resources.

The third stage centered on conservation and representational reconstruction. When the target shifts to engineering seepage, pointwise error alone is insufficient. Drawdown near wells, interlayer exchange, fracture-matrix flow, and the advance of multiphase fronts all depend on flux continuity and mass conservation. The weak-form theory-guided neural network (TgNN-wf) writes PDE residuals in integral form, uses integration by parts to reduce errors associated with high-order derivatives and automatic differentiation at discontinuities, and employs local test functions to capture local non-smoothness [12]. Mixed pressure-velocity PINNs treat pressure and velocity as primary variables simultaneously, avoiding the indirect recovery of fluxes through pressure derivatives. They are particularly suitable for high-contrast or discontinuous hydraulic-conductivity fields [14, 15].

At this stage, finite-volume residuals became another key route. The physics-informed convolutional neural network (PICNN) uses a convolutional neural network (CNN) to represent transient two-phase Darcy flow and constructs the loss from FVM discrete residuals, so that flux continuity between cells with different properties can be rigorously defined [16]. Building on this idea, E-PINN uses the embedded discrete fracture model (EDFM) for explicit fracture representation and combines FVM residuals, adjacency-location anchoring, adaptive activations, skip connections, and gated updates to handle two-dimensional and three-dimensional fractured multiphase flow [17].

Such methods show that traditional numerical schemes are not competitors to PINNs; rather, they provide the structural backbone for improving physical fidelity.

The fourth stage began to treat PINNs as differentiable modules within engineering workflows rather than emphasizing a single MLP that solves one fixed problem at a time. The finite-volume physics-informed Fourier neural operator (FV-PIFNO) combines a finite-volume physics loss with a Fourier neural operator, shifting the objective from single-instance solution to learning mappings from families of parameter fields to families of solution fields [18]. The 2024 work of Bandai et al. combines a nonlinear solver for Richards' equation with free-form constitutive learning through JAX and implicit differentiation [19]. Li et al. integrate PINNs with Hydrus-1D data, using the traditional model to provide low-cost priors while constraining infiltration processes through physics losses [20]. PINN work based on hydrogeophysical data uses electrical resistivity tomography (ERT) and point-sensor data for root-water-uptake and saturation reconstruction, showing its potential in monitoring interpretation [21]. Table 2 and Figure 2 summarize this methodological evolution and the seepage bottlenecks that motivated each stage.

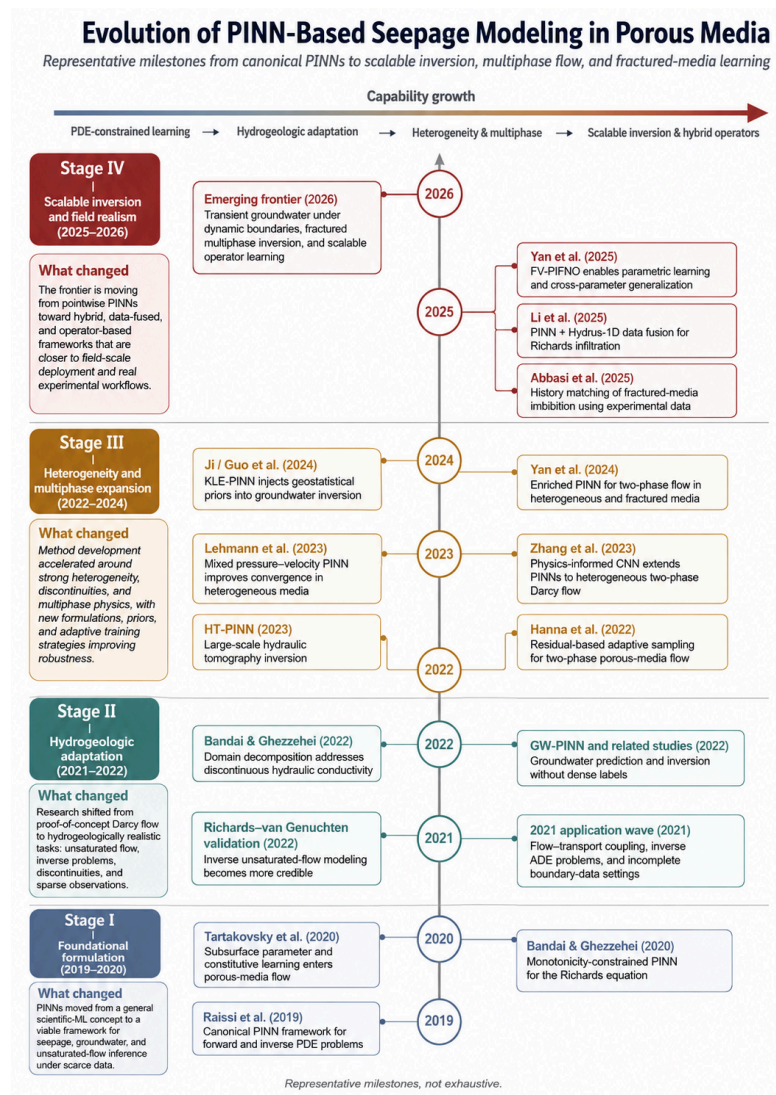


Figure 2. Evolution of PINN-based seepage modeling in porous media

Table 2. Method evolution and seepage bottlenecks

Stage	Methodological response	Main bottlenecks	Benefited tasks
I	Joint learning of parameters, states, and constitutive relations	Sparse observations, unobservable parameters, missing constitutive laws	Inverse problems, reconstruction
II	Monotonicity constraints, domain decomposition, adaptive sampling	Ill-posedness, non-smoothness, localized high gradients	Inverse problems, reconstruction, complex forward problems
III	Weak-form residuals, mixed variables, FVM-based losses	Insufficient conservation, flux discontinuity, multiphase fronts	Forward problems and high-fidelity inversion
IV	Fracture embedding, solver coupling, neural operators	Engineering geometries, real noise, parametric prediction	Synergy across all three task types

4. Task-oriented applications

Unlike the previous section, which tracks structural changes in PINNs through methodological evolution, this section organizes applications according to the "dominant unknown." If the governing equations, material parameters, and boundary conditions are largely known and the unknowns are mainly pressure, hydraulic head, saturation, or flux fields, the work is classified as forward problem solution. If observations cover only part of the state and the goal is to identify hydraulic conductivity, constitutive relations, boundary fluxes, source/sink terms, or root-water-uptake terms, the work is classified as inverse modeling. If parameters are not the sole focus and the objective is to complete sparse sensor, ERT, or point-cloud observations into continuous spatiotemporal fields, the work is classified as flow-field reconstruction. Many papers have attributes of two categories at once. For example, operator learning can serve as a forward surrogate and can also participate in inversion through data terms. The classification here is based on the dominant unknown and the most prominent technical contribution of each work. Table 3 summarizes the task-oriented classification used in the following subsections.

Table 3. Task-oriented summary of PINN applications in porous media seepage

Task type	Primary unknowns	Typical inputs	Expected outputs	Representative emphasis
Forward problem solution	State variables (e.g., hydraulic head, pressure, saturation, flux, temperature)	Known governing equations (PDEs), material parameters, boundary conditions, and initial conditions	Complete spatiotemporal field distributions	Surrogate solvers or mesh-free PDE solution
Inverse modeling	Parameters, constitutive relations, boundary terms, source terms, interface locations	Partial observation data + PDE constraints	Identified parameters or latent driving mechanisms	Ill-posed problem estimation under physical constraints
Flow-field reconstruction	Unobserved state fields in most of the domain	Sparse or incomplete measurement data + governing equations	Full-field reconstruction results satisfying physical constraints	Support for data assimilation and monitoring

This perspective also avoids reducing PINNs to the binary question of whether they can replace solvers. For forward problems, the key criteria are local conservation, flux continuity, interface

treatment, and computational efficiency. For inverse problems, the key criteria are identifiability, physical plausibility of prior constraints, noise robustness, and independent validation. For reconstruction tasks, the key criteria are extrapolation reliability in unobserved regions, mass closure, and uncertainty coverage. The following subsections discuss specific papers to clarify what problem PINNs actually solve in each task class, what assumptions they rely on, and what engineering risks remain.

4.1. Forward modeling: from solution approximation to conservative field representation

Forward problem solution is the most intuitive entry point for PINNs, but in seepage problems it also most readily exposes their disadvantages relative to FEM/FVM. When hydraulic parameters, boundary conditions, and initial conditions are all known, mature numerical solvers are usually stronger in speed, robustness, local conservation, and engineering auditability. Therefore, the academic value of forward PINNs should not be understood as "solving Darcy's equation once again with a neural network." Rather, it should be evaluated by how they explicitly encode the quantities that seepage engineering truly cares about - fluxes, interface conditions, strong gradients near sources and sinks, and fracture-matrix exchange - into the network representation. In other words, the central test for forward PINNs is not whether the pressure-field image is smooth, but whether the network can maintain an interpretable conservation structure under heterogeneous, discontinuous, and multiphase conditions.

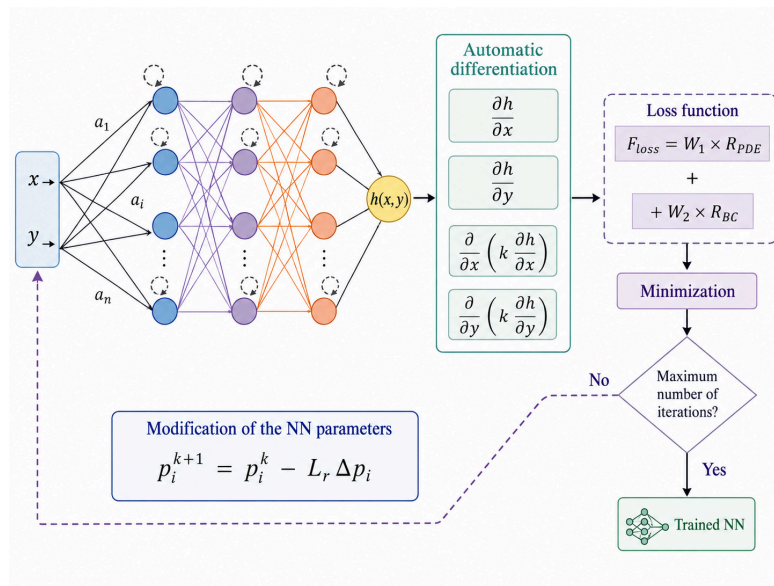


Figure 3. Strong-form PINN training workflow for seepage modeling

The mixed pressure-velocity PINN of Lehmann et al. is representative of this line of development. A traditional pressure-only PINN treats hydraulic head as the only unknown and then computes velocity or flux by automatic differentiation. This approach can work when hydraulic conductivity is continuous and smooth. However, in high-contrast or discontinuous media, the strong-form residual requires differentiating $K(x)\nabla h$, which is equivalent to asking the network to handle the non-differentiable structure of K . Lehmann et al. reformulated the governing equation as two sets of residuals, mass conservation and Darcy's law, and treated pressure (head) and velocity components as primary variables at the same time. In this way, training no longer requires spatial derivatives of discontinuous hydraulic conductivity; instead, $\nabla \cdot \mathbf{q} = 0$ and $\mathbf{q} + K\nabla h = 0$ are

constrained separately. The paper compared different network organizations and found that assigning separate networks to pressure and to each velocity component achieved the highest accuracy across several heterogeneous fields without substantially increasing training time. The key contribution of this work is not merely "adding output variables," but transforming the forward seepage problem from pressure regression into flux-consistent field representation [14]. A generic strong-form PINN training workflow is illustrated in Figure 3.

Alhubail et al. further showed that mixed variables do not automatically ensure stable training. The pressure-velocity formulation decomposes one PDE residual into multiple loss components, including the continuity equation, Darcy relation, and boundary terms. These terms differ in physical dimension, numerical scale, and gradient magnitude. If they are simply summed with equal weights, the optimizer may preferentially satisfy terms that decrease easily while neglecting constraints that are more critical for fluxes or boundaries. The paper systematically compared several global and local weighting strategies and tested their performance in random heterogeneous and blockwise discontinuous media. The conclusion was clear: max-average weighting improved convergence in random heterogeneous fields and was the only strategy that converged stably in discontinuous heterogeneous fields; under high heterogeneity, the unweighted baseline PINN became unstable. Thus, the bottleneck of forward seepage PINNs is not only network expressiveness, but also optimization balance among multiphysics residuals [15].

For transient two-phase Darcy flow, the PICNN proposed by Zhang et al. advances forward PINNs from continuous strong-form residuals to discrete conservative residuals. The paper studies heterogeneous oil-water two-phase flow with source/sink terms. It uses a convolutional network to represent the pressure field, but the physics loss is constructed not by automatic differentiation, but by approximating the PDE residual with a finite-volume scheme. This brings two direct benefits. First, flux continuity between grid cells with different properties can be rigorously defined at the level of discrete face fluxes. Second, high pressure gradients near wells can be handled separately with a well model, rather than forcing an ordinary MLP/CNN to "hard fit" sharp peaks by uniform sampling. PICNN also adopts the implicit-pressure, explicit-saturation (IMPES) idea, separating implicit pressure solution from explicit saturation updating; therefore, only one CNN needs to be trained at each time step. Dirichlet boundary conditions are not required, but they can be used as label-like constraints to improve accuracy. The importance of this paper lies in embedding the conservation backbone of traditional FVM into neural-network training, allowing the network to function both as a label-free solver and, when sufficient samples are available, as a surrogate model for two-phase flow [16].

E-PINN extends this discrete-conservation logic to fractured porous media. Yan et al. addressed two-phase displacement in two-dimensional and three-dimensional heterogeneous and fractured reservoirs. They used EDFM to explicitly represent fractures rather than expecting a global smooth network to learn fracture channels implicitly. At the same time, they constructed the physics loss with FVM so that flux continuity between matrix and fractures, between fractures, and between cells with different properties could be explicitly constrained. The network architecture also incorporates stronger physical inductive biases: adjacency-location anchoring encodes adjacency and location information; adaptive activations, skip connections, and gated updates enhance pressure-information propagation; and initial and boundary conditions are written into the network design through hard constraints. Because two-phase-flow training is difficult, E-PINN also uses IMPES, allowing the neural network to solve only for pressure while saturation is advanced explicitly. Relative to PICNN, the advance of E-PINN is that it extends "conservation in heterogeneous grids" into an integrated

framework of "explicit fracture-geometry embedding + discrete conservation + structure-enhanced architecture" [17].

As engineering needs shift from solving a single case to rapidly evaluating multiple scenarios, forward modeling has further moved toward operator learning. FV-PIFNO uses a Fourier neural operator to learn the mapping from heterogeneous parameter fields to solution fields, but still defines the physics loss with finite-volume discrete equations. This bypasses the difficulty of using automatic differentiation on discontinuous permeability fields while preserving flux continuity between adjacent cells. Its Gated-FNO/space-frequency filtering design attempts to exploit global frequency-domain dependence and local noise-suppression capability simultaneously [18]. In unsaturated infiltration, the physics-informed DeepONet of Kamil et al. learns an operator mapping from initial soil-water states and infiltration or drip-irrigation fluxes to full spatiotemporal profiles of pressure head or water content, and uses physics or limited data for extrapolation fine-tuning. The PI-DeepONet of Ye et al. further integrates varying boundary conditions, empirical constitutive relations, and limited observations into the same operator framework to simultaneously handle forward prediction, missing boundary-flux reconstruction, and constitutive-relation reconstruction [22, 23].

Another, more engineering-oriented route in forward problems is the coupling of PINNs with numerical solvers. Li et al. combined PINNs with Hydrus-1D data to solve the Richards infiltration equation. The core idea is not to discard mature solvers entirely, but to use a small amount of Hydrus-1D data to mitigate the sensitivity of PINNs to parameter initialization and complex boundaries, while using PDE constraints to maintain physical consistency in data-scarce regions. Such work indicates that, in strongly nonlinear unsaturated seepage with complex boundaries, the more realistic role of PINNs is often as a differentiable enhancement layer for numerical models rather than as stand-alone replacements [20].

Accordingly, representative progress in forward problem solution can be summarized in three steps. Mixed variables elevate fluxes from derivative-based byproducts of pressure to primary unknowns. FVM/EDFM residuals explicitly embed conservation and fracture geometry in the loss. Neural operators and solver coupling extend single-instance solution to rapid prediction across parameters and boundary conditions. Even so, for large-scale three-dimensional engineering forward modeling with fully known inputs, PINNs still have difficulty outperforming FEM/FVM overall. Their more valuable position is as differentiable solvers, local physics-completion modules, fast surrogate models, and front-end physics engines for inversion and reconstruction workflows.

4.2. Inverse modeling: from parameter fitting to identifiable physics

Inverse modeling is the task in which PINNs have the clearest comparative advantage in porous-media seepage. Field sites usually provide only sparse data on hydraulic head, water content, pressure, concentration, electrical resistivity, or production rates, whereas hydraulic conductivity, transmissivity, soil hydraulic functions, boundary fluxes, well source/sink strengths, and root-water-uptake terms are the key unknowns controlling prediction. Traditional inversion typically requires repeated forward modeling and regularization choices. The advantage of PINNs is that observation errors, governing equations, and structural priors can be written as a single differentiable objective function, thereby transforming a "lack of data" problem into an "identifiability under physical constraints" problem.

Early work by Tartakovsky et al. and He et al. established this idea. Tartakovsky et al. used physics-constrained DNNs to learn states, parameters, and unknown constitutive relations simultaneously, showing that even functions without direct observations can be identified under

state observations and PDE constraints [8]. He et al.'s MPINN further coupled subsurface flow with solute transport. Independent DNNs represented spatially variable hydraulic conductivity, hydraulic head, and concentration, and were trained jointly under residuals of Darcy's equation and the advection-dispersion equation. A key finding was that joint inversion of multiphysics variables can improve parameter-estimation accuracy. This indicates that the core of PINN inversion is not single-variable regression, but the use of complementary multi-field information to improve the identifiability of hidden parameters [9].

For unsaturated-flow inversion, Bandai and Ghezzehei's monotonicity-constrained RRE-PINN is among the most representative studies. This method uses time series of volumetric water content as the main observations and does not require complete initial and boundary conditions. The network is composed of three coupled feedforward networks, two of which represent the water retention curve (WRC) and the hydraulic conductivity function (HCF), respectively, and are explicitly constrained to be monotonic functions. The reason for this design is that Richards' equation alone cannot exclude all physically absurd constitutive curves. Only by writing the shape priors of the WRC/HCF into the network can the feasible inversion space be genuinely narrowed. The paper tested noise robustness using synthetic data generated by HYDRUS-1D for sandy loam, loam, and silt loam, and showed that water flux density can be derived from the trained PINN. Its limitations are also important for a review: errors are larger at the wetting front and in near-surface abrupt-transition regions, and high-quality inversion still depends on sufficiently dense water-content observations [10]. A representative RRE-PINN architecture for unsaturated-flow inversion is shown in Figure 4.

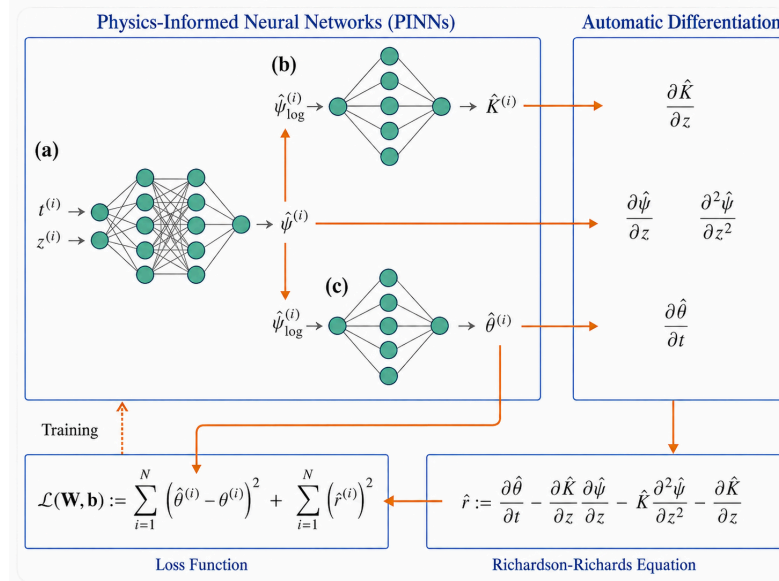


Figure 4. Representative RRE-PINN architecture for constitutive inversion

The physically constrained neural networks of Bandai et al. (2024) move constitutive inversion from low-dimensional empirical models toward free-form function learning. Traditional van Genuchten-Mualem-type models have only a few parameters, which facilitates calibration but may over-compress real soil hydraulic properties. This work constructs physically constrained neural networks in the differentiable JAX framework and obtains efficient gradients through implicit differentiation of a nonlinear Richards-equation solver, thereby learning more flexible WRC/HCF shapes. The authors tested different noise levels and network degrees of freedom on noisy synthetic data and upward-infiltration experimental data, and showed that neural-network constitutive

relations better fit the experimental data than empirical parameterized models. The significance is that PINN inversion is not limited to estimating a few empirical parameters; it can also release the constitutive-function space under physical constraints [19].

When the inversion target has layered structures or abrupt hydraulic-conductivity changes, the domain-decomposition PINN of Bandai and Ghezzehei provides another solution. For the one-dimensional RRE, this work showed that a PINN with adaptive activation functions can reach accuracy close to that of conventional numerical solutions in homogeneous soils. For layered soils, separate networks are assigned to individual soil layers and connected through interface conditions, so that water dynamics are no longer forcibly smoothed by a single network. More importantly, the framework can estimate unspecified upper or lower boundary water fluxes from sparse soil-moisture observations and outputs the domain-wide water flux as a byproduct. This is important for engineering monitoring because boundary fluxes are often more difficult to measure directly than internal water content [11].

For high-dimensional parameter inversion in saturated groundwater systems, hydraulic tomography PINN (HT-PINN) and Karhunen-Loeve expansion PINN (KLE-PINN) demonstrate two different routes. The HT-PINN of Guo et al. targets hydraulic tomography, representing the transmissivity field by one neural network while using a series of networks to describe multiple steady-state or transient sequential pumping tests. All networks are trained jointly through hydraulic-head data errors and PDE residuals. Owing to collocation-point batch training, this method can invert a two-dimensional Gaussian transmissivity field with a resolution of 1024×1024 and more than one million unknowns; the authors reported inversion-image accuracy exceeding 95% [24]. The KLE-PINN of Ji et al. represents the hydraulic-conductivity field using a Karhunen-Loeve expansion, reducing parameter dimensionality with prior geostatistical information. They compared three data scenarios - single-well pumping, natural gradient, and hydraulic tomography - and found that multiple pumping-test data led to more accurate estimates, faster convergence, and higher stability. The framework is also convenient for studying cases with unknown outer boundary conditions [25].

Recent inversion targets have expanded from hydraulic parameters to ecohydrological source/sink terms and observation operators. Sakar et al. constructed a two-output PINN for hydrogeophysical data, using ERT and sparse point-sensor data to reconstruct both a high-resolution saturation field and an unknown root-water-uptake (RWU) function. When an additional measurable constraint on daily total transpiration was introduced, the root mean square error (RMSE) of daily transpiration decreased by approximately 88%, local RWU error could be reduced to below 5%, and Feddes water-stress response parameters could be recovered [21]. Such work advances PINNs from "hydraulic-parameter inverters" to joint interpretation tools linking monitoring data, state fields, and biophysical source/sink terms. In an adjacent direction, coupling Hydrus-1D with PINNs also shows that a small amount of numerical-model data can improve the robustness of Richards-equation inversion or solution under complex boundaries and initialization [20].

History matching in fractured porous media pushes inverse PINNs toward scenarios closer to experimental data. Abbasi et al. constructed a PINN workflow for multiphase imbibition in naturally fractured shale core plugs. Experimentally, the core was first saturated with CO_2 and then injected with brine, while computed tomography (CT) scanning was used to obtain fracture locations and the dynamic in situ distribution of fluids. In modeling, the matrix and fracture domains were decomposed into different subnetworks, pressure and saturation were represented separately, and multiphase-flow governing equations, interfacial exchange, and observations were jointly constrained. Because fracture-matrix interactions significantly alter effective relative permeability

and capillary behavior, the paper also used a pretraining strategy to reduce the optimization difficulty of direct joint inversion and discussed parameter uniqueness and uncertainty. Its significance is that PINN inversion is no longer limited to verifying hydraulic conductivity or constitutive parameters against synthetic truth. It is beginning to enter a history-matching workflow of "experimental observation - geometry identification - multi-network inversion - uncertainty checking." Its limitations are equally clear: fracture geometry still requires high-quality priors such as CT data, and core-scale conclusions still require additional validation before being extended to field scale [26].

Nevertheless, the success of inverse PINNs must be interpreted cautiously. PDE residuals and neural-network priors can compress the feasible solution space, but they cannot mathematically eliminate non-uniqueness. If observation locations are insensitive, boundary conditions are missing, the noise model is misspecified, or the prior structure is too strong, the network may still return parameter fields that look smooth but are not identifiable. Therefore, reviews of inverse-problem PINNs should prioritize four kinds of evidence: whether independent observations or experimental validation are provided; whether the influence of observation number and location is analyzed; whether noise robustness, prior sensitivity, and uncertainty are reported; and whether the inferred parameters improve extrapolative prediction rather than merely fitting the training interval.

4.3. Flow-field reconstruction: physics-constrained completion of sparse observations

Flow-field reconstruction lies between forward modeling and inversion. Researchers do not necessarily require exact identification of every material parameter, but they need to recover continuous spatiotemporal state fields from sparse, incomplete, or indirect observations, such as hydraulic head, pressure, saturation, water content, pore-scale velocity, or root-zone moisture distribution. Compared with pure interpolation or data-driven completion, the advantage of PINNs is that PDE residuals constrain unobserved regions, so that reconstructed fields are not merely visually smooth but also satisfy, as far as possible, mass conservation, boundary conditions, and source/sink mechanisms.

The GW-PINN of Zhang et al. is a core contribution in groundwater reconstruction. This method does not rely on labeled solutions; instead, it constructs the loss directly from the groundwater-flow PDE, initial and boundary conditions, and hard/soft constraints. To address sharp hydraulic-head changes near wells or source/sink terms, it proposes locally refined spatial-temporal sampling, concentrating collocation points in the local regions where resolution is truly needed. To address training difficulty over long time domains, it also proposes snowball-style two-stage training, splitting the time domain and advancing progressively. Five examples cover scenarios such as confined and unconfined aquifers with pumping wells, multiple pumping wells, and two parallel channels, and the results are compared with MODFLOW or analytical solutions. The results show that the hard-constraint version is more robust than the soft-constraint version, and that locally refined sampling achieves higher accuracy with fewer training points. The framework can also serve as a surrogate model for parameterized groundwater-flow equations [13]. Figure 5 illustrates the network representation and sampling strategies relevant to this type of groundwater reconstruction.

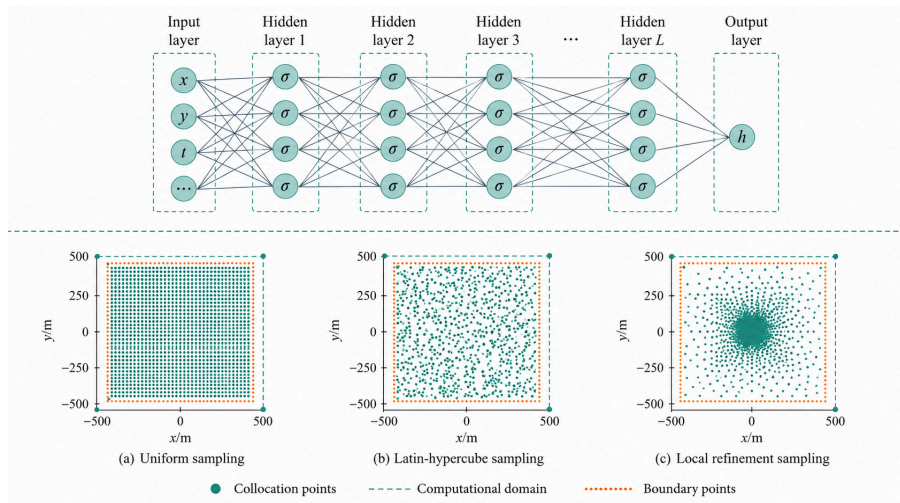


Figure 5. Network representation and collocation sampling strategies in groundwater PINNs

Secci et al. advanced reconstruction to transient unconfined aquifers. The difficulty in unconfined aquifers is that the phreatic surface itself is a boundary that varies in time and space, rather than an ordinary Dirichlet or Neumann condition on a fixed geometry. The paper estimates piezometric head and phreatic-surface evolution in a vertical two-dimensional cross section, examining both a homogeneous isotropic aquifer and a mildly heterogeneous anisotropic aquifer. It emphasizes the potential of PINNs for handling time-varying free boundaries in data-scarce environments. Its complementarity with GW-PINN lies in the fact that GW-PINN emphasizes local refinement near well/source-sink terms and time decomposition, whereas Secci et al. emphasize the challenge posed to reconstruction tasks by a dynamic phreatic surface as a moving boundary [27].

Virupaksha et al. further tested unconfined-aquifer reconstruction under dynamic boundaries and time-varying pumping/drainage conditions. The paper notes that many existing groundwater PINNs remain focused on confined aquifers or steady-state settings, whereas the governing equations for unconfined aquifers are more nonlinear and require more collocation points for training. The authors therefore focus on how the time derivative is handled and systematically compare continuous-time PINNs, discrete-time PINNs, and time-decomposition PINNs. Multiple tests with variable boundary conditions or pumping rates show that discrete-time methods outperform standard continuous-time PINNs in both accuracy and training efficiency, and can reduce training time by about one order of magnitude by decreasing the number of trainable parameters, training epochs, and collocation points. For state-field reconstruction, this result indicates that, when the problem enters long-duration dynamic processes, sampling refinement is only the first layer of strategy; time discretization and subproblem organization are equally decisive for stable recovery of spatiotemporal hydraulic-head fields [28].

Pore-scale flow-field reconstruction faces entirely different difficulties: the geometry is not a simple rectangular domain, but an irregular connected space formed by pore-grain boundaries. The physics-informed PointNet of Kashеfi and Mukerji uses pore-space point clouds as input to predict steady Stokes-flow pressure-velocity fields and trains with sparse observations and physics residuals. Compared with CNNs based on images or regular grids, PointNet has the advantage of inputting only the pore space rather than voxelizing both solid grains and pores, thereby reducing memory demand. Pore boundaries can be represented more smoothly, and spatial resolution can vary with geometric complexity. The paper also examines the effects of noisy sensor data, pressure

observations, and spatial correlation length, making the study closer to reconstruction problems in real pore-scale measurement and digital rock analysis [29].

For unsaturated-infiltration reconstruction, the physics-informed deep learning (PIDL) framework of Lan et al. combines sparse data, low-/high-fidelity data fusion, and uncertainty quantification. The method has three components: a PINN uses Richards' equation and a small amount of noisy observation data to recover infiltration processes; when high-fidelity observations are too scarce, a multi-fidelity PINN (MF-PINN) introduces more readily available low-fidelity data to assist reconstruction; and Monte Carlo dropout (MC-dropout) provides prediction uncertainty intervals. The paper also allows unknown parameters in the soil-water characteristic curve to be identified while reconstructing the water-content field, and validates the framework using three unsaturated-infiltration models and one field drainage case. Its significance is that it advances the task from "providing one reconstructed field" to "providing a physically consistent reconstructed field with uncertainty intervals," which is precisely the form needed for engineering monitoring [30].

Hydrogeophysical PINNs further show that reconstruction tasks often require indirect observations to be converted into state fields. Sakar et al. do not use ERT images directly as the final result. Instead, they treat ERT and point-sensor data as constraints and use Richards' equation to connect them with the saturation field and root-water-uptake terms. The framework therefore estimates the spatiotemporal distribution of RWU while reconstructing high-resolution saturation, giving it both state-reconstruction and inverse-problem attributes [21]. Operator learning is also changing the scale of reconstruction tasks. Once trained, physics-informed DeepONet and PI-DeepONet can rapidly output complete spatiotemporal profiles under varying initial conditions, boundary fluxes, or soil configurations, thereby advancing single reconstruction toward real-time completion across multiple scenarios [22, 23].

Overall, flow-field reconstruction may be one of the most natural current engineering landing points for PINNs, because it uses the continuous-field representation of neural networks while avoiding the least favorable competition with mature solvers in standard forward problems with fully known inputs. However, evaluation criteria must be strict. A reconstructed field should not be assessed only by whether it is smooth or matches a few observation points. Reports should also include errors at sensors not used for training, global PDE residuals, volume/mass closure errors, errors in key local regions such as well vicinities, wetting fronts, and fracture intersections, and the coverage rate of uncertainty intervals. Only when these metrics are satisfied can PINN reconstruction move from visually appealing results to engineering information that can support monitoring, early warning, and decision-making.

5. Challenges and outlook

First, there remains a mismatch between expressive capacity and non-smooth physics. Standard MLPs have a low-frequency bias and struggle to capture fracture channeling, multiphase fronts, and sharp gradients near wells. Fourier features, Kolmogorov-Arnold networks (KANs), PointNet, graph neural networks, and neural operators offer new routes for multiscale representation, but their advantages in seepage must be tested on unified benchmarks rather than generalized from a small number of examples [18, 29, 31-33].

Second, soft-constraint residuals are insufficient to guarantee engineering conservation. Seepage engineering cares about mass closure, flux continuity, and strict satisfaction of boundary conditions, whereas ordinary PINN loss minimization can only provide approximate satisfaction. More robust directions include mixed variables, weak-form/variational residuals, FVM/FEM discrete residuals,

hard boundary constructions, and PINN-solver coupling. Neural networks should undertake differentiable inversion, surrogate prediction, and data fusion; traditional numerical schemes should provide the backbone for conservation and geometric handling [12, 14, 16-18].

Third, training optimization remains an ill-conditioned multi-objective problem. PDE residuals, boundary terms, initial terms, data terms, and prior terms differ in physical dimension, scale, and gradient direction. Fixed weights can easily cause some constraints to be ignored. Adaptive weighting, neural-tangent-kernel (NTK)-inspired loss balancing, sequential time training, residual/failure-informed sampling, and domain decomposition will become necessary components [13, 15, 34, 35]. For seepage, collocation strategies should serve the physical difficulties of the problem rather than pursue uniform full-domain coverage.

Fourth, uncertainty quantification and identifiability analysis remain insufficient. Many studies report only mean errors without explaining whether parameters are unique, how observation noise propagates, or whether model extrapolation is reliable. Future work should combine Bayesian PINNs, ensemble learning, conformal prediction, generalized fiducial inference, or Gaussian-process calibration to provide confidence intervals and risk bounds. Deterministic reconstruction without UQ is unlikely to meet the requirements of high-risk applications such as dam safety, contaminant source tracing, or CO₂ storage.

Fifth, domain benchmarks and real-data validation need to be strengthened. Existing studies often use self-generated numerical solutions as ground truth, which can easily overestimate generalization capability. More valuable benchmarks should include heterogeneity, fractures, noisy observations, missing boundaries, three-dimensional geometries, and independent experimental data, and should report errors, conservation residuals, computational cost, and uncertainty coverage in a unified way. Only under such standards can PINNs move from proof of concept into reliable engineering toolchains.

6. Conclusion

PINNs provide a modeling approach for seepage in porous media that differs from traditional discrete solvers. They do not first discretize the governing equations; instead, they place state fields, parameters, constitutive relations, and observational data within a single differentiable optimization framework. For this reason, their most prominent value is not replacing FEM/FVM in standard forward problems, but handling sparse-data inversion, unknown constitutive learning, and physics-constrained flow-field reconstruction where traditional workflows are difficult.

From a historical perspective, the development of seepage PINNs has always been driven by specific physical bottlenecks. Early work demonstrated that PDE constraints can support joint parameter-state-constitutive learning. Monotonicity constraints and domain decomposition improved the credibility of unsaturated constitutive inversion and discontinuous-medium modeling. Weak-form residuals, mixed pressure-velocity variables, and finite-volume losses shifted attention toward fluxes and local conservation. E-PINN, FV-PIFNO, and solver-coupling frameworks further show that PINNs are moving from stand-alone networks toward differentiable modules in engineering workflows.

From the perspective of task maturity, the three application classes are uneven. Forward problem solution remains limited by training cost, conservation consistency, and three-dimensional scalability, and is more suitable as a surrogate, local solver, or inversion module. Inverse modeling best reflects the strength of PINNs in fusing observations, equations, and priors, but it must be accompanied by identifiability and uncertainty analysis. Flow-field reconstruction is currently the

most natural engineering landing point, especially in scenarios with sparse monitoring, incomplete boundaries, and a need for continuous spatiotemporal interpretation.

Future reliable seepage PINNs should not pursue the goal of "pure neural networks replacing all numerical models." Instead, they should move toward hybridization, conservation, and uncertainty awareness. Finite-volume/finite-element structures should be responsible for conservation and complex geometry, while neural networks should be responsible for differentiable data assimilation, unknown closure learning, and fast cross-scenario prediction. Only when algorithm validation expands from low-dimensional synthetic examples to real monitoring, fractured media, multiphase processes, and high-risk engineering decisions can PINNs become trustworthy, interpretable, and deployable computational tools for simulating seepage in porous media.

References

- [1] Bear, J. Dynamics of Fluids in Porous Media. Elsevier, 1972.
- [2] Richards, L. A. Capillary conduction of liquids through porous mediums. *Physics*, 1931, 1, 318–333.
- [3] Eymard, R., Gallouët, T., and Herbin, R. Finite volume methods. *Handbook of Numerical Analysis*, 2000, 7, 713–1018.
- [4] Raissi, M., Perdikaris, P., and Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 2019, 378, 686–707.
- [5] Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., and Yang, L. Physics-informed machine learning. *Nature Reviews Physics*, 2021, 3, 422–440.
- [6] Cuomo, S., Schiano di Cola, V., Giampaolo, F., Rozza, G., Raissi, M., and Piccialli, F. Scientific machine learning through physics-informed neural networks: where we are and what's next. *Journal of Scientific Computing*, 2022, 92, 88.
- [7] Yang, G., Xu, R., Tian, Y., Guo, S., Wu, J., and Chu, X. Data-driven methods for flow and transport in porous media: a review. *International Journal of Heat and Mass Transfer*, 2024, 235, 126149.
- [8] Tartakovsky, A. M., Marrero, C. O., Perdikaris, P., Tartakovsky, G. D., and Barajas-Solano, D. Physics-informed deep neural networks for learning parameters and constitutive relationships in subsurface flow problems. *Water Resources Research*, 2020, 56, e2019WR026731.
- [9] He, Q. Z., Barajas-Solano, D., Tartakovsky, G., and Tartakovsky, A. M. Physics-informed neural networks for multiphysics data assimilation with application to subsurface transport. *Advances in Water Resources*, 2020, 141, 103610.
- [10] Bandai, T., and Ghezzehei, T. A. Physics-informed neural networks with monotonicity constraints for Richardson-Richards equation: estimation of constitutive relationships and soil water flux density from volumetric water content measurements. *Water Resources Research*, 2021, 57, e2020WR027642.
- [11] Bandai, T., and Ghezzehei, T. A. Forward and inverse modeling of water flow in unsaturated soils with discontinuous hydraulic conductivities using physics-informed neural networks with domain decomposition. *Hydrology and Earth System Sciences*, 2022, 26, 4469–4495.
- [12] Xu, R., Zhang, D., Rong, M., and Wang, N. Weak form theory-guided neural network (TgNN-wf) for deep learning of subsurface single- and two-phase flow. *Journal of Computational Physics*, 2021, 436, 110318.
- [13] Zhang, X., Zhu, Y., Wang, J., Ju, L., Qian, Y., Ye, M., and Yang, J. GW-PINN: A deep learning algorithm for solving groundwater flow equations. *Advances in Water Resources*, 2022, 165, 104243.
- [14] Lehmann, F., Fahs, M., Alhubail, A., and Hoteit, H. A mixed pressure-velocity formulation to model flow in heterogeneous porous media with physics-informed neural networks. *Advances in Water Resources*, 2023, 181, 104564.
- [15] Alhubail, A., Fahs, M., Lehmann, F., and Hoteit, H. Modeling fluid flow in heterogeneous porous media with physics-informed neural networks: weighting strategies for the mixed pressure head-velocity formulation. *Advances in Water Resources*, 2024, 193, 104797.
- [16] Zhang, Z., Yan, X., Liu, P., Zhang, K., Han, R., and Wang, S. A physics-informed convolutional neural network for the simulation and prediction of two-phase Darcy flows in heterogeneous porous media. *Journal of Computational Physics*, 2023, 477, 111919.

- [17] Yan, X., Lin, J., Wang, S., Zhang, Z., Liu, P., Sun, S., Yao, J., and Zhang, K. Physics-informed neural network simulation of two-phase flow in heterogeneous and fractured porous media. *Advances in Water Resources*, 2024, 189, 104731.
- [18] Yan, X., Lin, J., Ju, Y., et al. A finite-volume based physics-informed Fourier neural operator network for parametric learning of subsurface flow. *Advances in Water Resources*, 2025, 205, 105087.
- [19] Bandai, T., Ghezzehei, T. A., Jiang, P., Kidger, P., Chen, X., and Steefel, C. I. Learning constitutive relations from soil moisture data via physically constrained neural networks. *Water Resources Research*, 2024, 60, e2024WR037318.
- [20] Li, Y., Sun, Q., Fu, Y., and Wei, J. Solving the Richards infiltration equation by coupling physics-informed neural networks with Hydrus-1D. *Scientific Reports*, 2025, 15, 18649.
- [21] Sakar, C., Tsukanov, K., Schwartz, N., and Moreno, Z. A physics-informed neural network workflow for forward and inverse modeling of unsaturated flow and root water uptake from hydrogeophysical data. *Journal of Hydrology*, 2026, 665, 134675.
- [22] Kamil, H., Soulaïmani, A., and Beljadid, A. Physics-informed neural operators for efficient modeling of infiltration in porous media. *Journal of Computational Physics*, 2025, 538, 114156.
- [23] Ye, Q., Huang, Z., Zheng, Q., and Zeng, L. Predicting water movement in unsaturated soil using physics-informed deep operator networks. *Advances in Water Resources*, 2025, 202, 105001.
- [24] Guo, Q., Zhao, Y., Lu, C., and Luo, J. High-dimensional inverse modeling of hydraulic tomography by physics informed neural network (HT-PINN). *Journal of Hydrology*, 2023, 616, 128828.
- [25] Ji, Y., Zha, Y., Yeh, T.-C. J., Shi, L., and Wang, Y. Groundwater inverse modeling: Physics-informed neural network with disentangled constraints and errors. *Journal of Hydrology*, 2024, 640, 131703.
- [26] Abbasi, J., Moseley, B., Kurotori, T., Jagtap, A. D., Kovscek, A. R., Hiorth, A., and Andersen, P. Ø. History-matching of imbibition flow in fractured porous media using physics-informed neural networks (PINNs). *Computer Methods in Applied Mechanics and Engineering*, 2025, 437, 117784.
- [27] Secci, D., Godoy, V. A., and Gómez-Hernández, J. J. Physics-informed neural networks for solving transient unconfined groundwater flow. *Computers & Geosciences*, 2024, 182, 105494.
- [28] Virupaksha, A., Lehmann, F., Hoteit, H., Younes, A., Fahs, M., and Nagel, T. Modeling transient groundwater flow in unconfined aquifers under dynamic conditions using physics-informed neural networks. *Water Resources Research*, 2026, 62(3), e2025WR040754.
- [29] Kashefi, A., and Mukerji, T. Prediction of fluid flow in porous media by sparse observations and physics-informed PointNet. *Neural Networks*, 2023, 167, 80–91.
- [30] Lan, P., Su, J., Zhu, S., Huang, J., and Zhang, S. Reconstructing unsaturated infiltration behavior with sparse data via physics-informed deep learning. *Computers and Geotechnics*, 2024, 168, 106162.
- [31] Wang, S., Wang, H., and Perdikaris, P. On the eigenvector bias of Fourier feature networks: from regression to solving multi-scale PDEs with physics-informed neural networks. *Computer Methods in Applied Mechanics and Engineering*, 2021, 384, 113938.
- [32] Liu, Z., Wang, Y., Vaidya, S., et al. KAN: Kolmogorov-Arnold Networks. *OpenReview/ICLR*, 2025.
- [33] Zhang, Z., Wang, Q., Zhang, Y., Shen, T., and Zhang, W. Physics-informed neural networks with hybrid Kolmogorov-Arnold network and augmented Lagrangian function for solving partial differential equations. *Scientific Reports*, 2025, 15, 9764.
- [34] Wang, S., Teng, Y., and Perdikaris, P. Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM Journal on Scientific Computing*, 2021, 43, A3055–A3081.
- [35] Gao, Z., Yan, L., and Zhou, T. Failure-informed adaptive sampling for PINNs. *SIAM Journal on Scientific Computing*, 2023, 45, A1971–A1994.