GROUNDWATER SEEPAGE MODELING IN A RIVER CANAL SYSTEM BASED ON PHYSICS-INFORMED NEU RAL NETWORKS

Anonymous authors

006

008 009 010

011

013

014

015

016

017

018

019

021

024

025

026

027

028

029

031

032

Paper under double-blind review

ABSTRACT

Neural networks, especially deep learning, have achieved revolutionary advances in several domains, including image and speech recognition, with excellent results. However, their reliance on labeled data, lack of interpretability, and inconsistency with physical principles limit their applicability in groundwater seepage prediction and other scientific disciplines. Physics-Informed Neural Networks (PINNs) significantly improve these issues by integrating physical knowledge with neural networks. This study focuses on modeling the groundwater flow field and proposes a physics-informed river-canal groundwater seepage model (PI-RGSM). This model enables self-supervised learning by incorporating hard constraints of boundary and initial conditions, utilizing hydrogeological parameters and boundary conditions as direct inputs, thus diminishing dependence on observable data. Compared to the baseline PINNs, the PI-RGSM adapts to and accurately predicts diverse seepage situations with just one training session, achieving a mean coefficient of determination of 0.978. To further enhance applicability in complex dynamic groundwater seepage situations, we propose PI-RGSM-K, which builds upon PI-RGSM. This model simulates heterogeneous groundwater seepage fields and improves performance in complex seepage environments through parameterized hydraulic conductivity field K(x, y) and fine-adjusted model architecture, attaining a mean coefficient of determination of 0.982. The physics-informed neural network models proposed in this study demonstrate exceptional efficacy in precisely forecasting groundwater seepage behavior.

033 1 INTRODUCTION

Groundwater is a crucial source of fresh water for the global population, essential for residential 035 life, agricultural irrigation, industrial production, and ecosystem protection (Massuel et al., 2018; Xie et al., 2023; Mishra, 2023). It is imperative to manage this rich resource sustainably to en-037 sure its long-term availability. Therefore, it is very important to understand groundwater dynamics. Field-based observations, theoretical analysis, hydrogeological modeling are essential ways to study groundwater dynamics (Li et al., 2021; Su et al., 2023). However, although some field-based obser-040 vations (e.g., pump tests, tracer Tests, monitoring wells, etc.) are inevitable, large amount of field-041 based observations for analyzing groundwater dynamics are time-consuming and labor-intensive; 042 although theoretical analysis is effective in providing deep understanding of the basic principles and 043 flexibility in exploring simplified complex systems, the oversimplification, heavy dependence on as-044 sumptions and lack of specificity limit the further application. Hydrogeological modeling is widely used to simulate groundwater dynamics for its relatively easy and cheap of implementing, ability of complex system representation, scenario analyzing, results visualization (Chen et al., 2020; Jing 046 et al., 2023). 047

With the evolution of scientific paradigms (Caíno-Lores et al., 2020), researches on natural science
has mitigated from empirical and theoretical approaches to computational and data-intensive scientific paradigms driven by computer simulations and big data (Babovi & Bajat, 2023; Li et al., 2023).
In the field of computational fluid dynamics, numerical models have achieved accurate simulations
of physical processes (refer to Appendix A). Important issues in those spatial discretization methods for solving the partial differential equations (PDEs) of groundwater are the computationally intensive. Continuous advancements in data-intensive scientific paradigms have empowered deep

054 neural networks (DNN) to achieve significant breakthroughs in many fields (refer to Appendix A). However, the "black box" nature of DNN exhibiting a lack of transparency in their decision-making 056 processes and the significant dependence on extensive training data, limit their use in groundwa-057 ter research. The physics-informed neural networks (PINNs) which combine the physical laws and 058 neural networks have improved the interpretability of models and diminished reliance on extensive datasets. Wang et al. (2020) proposed a theoretically guided neural network model (TgNN) based on the loss function of groundwater seepage differential equations, which improved the generalization 060 performance of the groundwater seepage model with limited observational data. Additionally, the 061 team demonstrated that the physics-informed model could effectively use physical information to 062 predict groundwater seepage responses beyond the training set through groundwater seepage trans-063 fer learning tasks. Wang et al. (2021a) proposed a neural network constrained by geostatistical infor-064 mation and a physics-guided autoencoder based on convolutional neural networks (CNNs) (Wang 065 et al., 2021b), applying the physical guidance method to the parameter inversion of groundwater 066 hydraulic conductivity fields. The model's accuracy and practicality were further enhanced. The 067 study by (Pashaei Kalajahi et al., 2022) demonstrated that, even with limited data, physics-guided 068 neural networks could accurately estimate key parameters such as hydraulic conductivity and poros-069 ity of the medium, showcasing their strong ability to combine data-driven approaches with physical guidance. Daolun et al. (2021) established a new specialized neuron model incorporating pressure gradient information and proposed an algorithm called signpost neural network (SNN), which signif-071 icantly improves the accuracy of solving unsteady seepage partial differential equations. Although 072 significant progress has been made in improving groundwater seepage models using PINNs, current 073 models still heavily depend on the adequacy and quality of observed data. Moreover, after the model 074 has been trained, its applicability is generally limited to specific hydrogeological parameter settings, 075 making it difficult to generalize to broader or unforeseen hydrogeological scenarios, thus restricting 076 the model's versatility and practicality. 077

Contributions: In this paper, a self-supervised deep learning method for groundwater seepage pre-078 diction, PI-RGSM, is proposed to address the limitations of baseline PINNs, which depends on ob-079 served data and have limited flexibility. Strategies of incorporating hard constraints combined with boundary condition inputs into the baseline PINNs and utilizing hydrogeological parameters and 081 boundary conditions as direct inputs are adopted. These allow the model to adapt new hydrogeolog-082 ical environments without retraining. Additionally, an extended version of the model, PI-RGSM-K, 083 is developed to simulate more complex heterogeneous seepage scenarios by parameterizing hy-084 draulic conductivity field K(x,y). This enhances the precision and adaptability of groundwater 085 seepage prediction, effectively overcoming the limitations of baseline models in heterogeneous seepage fields. Several representative river-canal groundwater seepage cases are designed to evaluate the accuracy and efficiency of the proposed method. 087

Paper Organization: The theoretical background is summarized in Section 2, along with a discussion of the improvement method for baseline PINNs. Section 3 introduces the construction and experimental setup of the PI-RGSM and PI-RGSM-K models. A series of experiments verifying the predictive capability of PI-RGSM and its extended model, PI-RGSM-K, under various hydrogeological conditions (including variations in precipitation, hydraulic conductivity, boundary conditions, and heterogeneous hydraulic conductivity field) are presented in Section 4. Finally, the summary and conclusions are provided in Section 5.

2 Method

095

096

097 2.1 GROUNDWATER SEEPAGE IN A RIVER-CANAL SYSTEM

098 As a key hydraulic engineering structure, river-canal system is a typical system which combines 099 river, canal and groundwater. The river-canal system significantly influences the surrounding ground-100 water flow, making its study crucial for groundwater resource assessment, artificial drainage system 101 design, and agricultural irrigation strategy formulation. Its uniqueness lies in considering scenarios 102 where two parallel river-canals cross the aquifers to reach impermeable bottom layers. As illustrated 103 in Figure 1, the groundwater between the two river-canals forms a specific flow pattern, with the left 104 and right canals serving as Dirichlet boundaries to create a convex hydraulic head surface, known as 105 the phreatic surface. The morphology of the phreatic line, represented by the cross-sectional curve of the phreatic hydraulic head H, is affected by the water levels H_a and H_b on both sides, as well 106 as rainfall infiltration W. In a phreatic aquifer, assuming negligible variation in hydrogeological 107 conditions in the vertical direction, groundwater flow is governed by the two-dimensional (in the

horizontal plane XY) seepage PDE (Equation 1). Where t is time (T); K is the hydraulic conductivitie (LT^{-1}) along the x, y axes, respectively. μ_s is the specific storage (L^{-1}) of the aquifer below the free surface. Full details in Appendix B.



Н

H

122 123 124

125

132

137

138 139 140

111

112

113 114

119 120

121

Figure 1: The schematic diagram of groundwater seepage in a river-canal system.

Impermeable bottom layers

126 2.2 LOSS FUNCTION OF BASELINE PINNS

In the equations of physics-informed groundwater seepage, the loss function is integrated with the basic PDE of groundwater seepage, initial and boundary conditions, observational data, and engineering control to guide the optimization direction of the neural network and ensure its physical consistency.

(1) **Loss function of PDE:** Considering the specific case of a two-dimensional heterogeneous seepage field with unit formation thickness, the seepage equation is:

$$\frac{\partial}{\partial x} \left(K(x,y) H \frac{\partial H(x,y,t)}{\partial x} \right) + \frac{\partial}{\partial y} \left(K(x,y) H \frac{\partial H(x,y,t)}{\partial y} \right) + W = \mu_s \frac{\partial H(x,y,t)}{\partial t}$$
(2)

Where K(x, y) is the hydraulic conductivity. Using this equation, we define the PDE residuals RES_{PDE} :

$$RES_{PDE} = \frac{\partial}{\partial x} \left(K(x,y)\hat{H}\frac{\partial\hat{H}}{\partial x} \right) + \frac{\partial}{\partial y} \left(K(x,y)\hat{H}\frac{\partial\hat{H}}{\partial y} \right) + W - \mu_s \frac{\partial\hat{H}}{\partial t}$$
(3)

141 When the physical consistency between the input and output variables of the model is higher, 142 RES_{PDE} approaches 0. During the calculation process, RES_{PDE} requires the model output (hy-143 draulic head \hat{H}) to perform higher-order differentiation with respect to the input spatial-temporal 144 coordinate points (x, y, t). This task can be performed directly and efficiently using the automatic 145 differentiation provided by modern machine learning frameworks such as TensorFlow or PyTorch. 146 The hyperbolic tangent function (Tanh) is chosen as the activation function in this study because its 147 favorable second-order differentiability meets the model's requirements.

(2) Constraint from definite solution condition: Consider the first type boundary condition
 (Dirichlet condition), where the water level at any given time at every point in a specific boundary region is known, which can be expressed as:

$$H(x,y,t)\big|_{\Gamma_1} = \varphi(x,y,t), \quad (x,y,t) \in \Gamma_1 \tag{4}$$

¹⁵² Where H(x, y, t) represents the hydraulic head at boundary Γ_1 , and $\phi(x, y, t)$ represents the hydraulic head function with spatial-temporal coordinates as the independent variable. By moving the function $\phi(x, y, t)$ to the left-hand side of the equation and substitute the hydraulic head on the boundary by the outputs of neural network (f_{net}) , the residual of boundary condition RES_{BC} is:

$$RES_{BC} = \hat{H} - \varphi(x, y, t), \quad (x, y, t) \in \Gamma_1$$
(5)

For initial conditions, this can be expressed as $\phi(x, y, 0) = H_0(x, y)$, where H_0 is the hydraulic head distribution function at the initial time. Therefore, the residual of initial condition RES_{IC} can be expressed as:

$$RES_{IC} = \varphi(x, y, 0) - \hat{H}, \quad t = 0 \tag{6}$$

(3) Constraint from engineering control: It is often necessary to incorporate specific ecological

156 157

161

167 168 169

170

171

172 173

174

175 176

177 178

201 202

208

209

and policy-oriented engineering control measures when solving actual groundwater seepage problems. For example, in ecological restoration, it may be essential to limit the decline of the hydraulic head in an area to a predetermined threshold H_{EC} . By constructing a residual term (RES_{EC}) for engineering control, it is possible to measure the discrepancy between the model prediction and the actual control objective:

$$RES_{EC} = \begin{cases} H_{EC} - \hat{H}, & H_{EC} > \hat{H} \\ 0, & H_{EC} \le \hat{H} \end{cases}$$
(7)

(4) **Constraint from observed data:** The difference between the hydraulic head predicted by the neural network and the actual hydraulic head observed can be quantified by constructing the residual RES_{OB} of the observed data:

$$RES_{OB} = H_{OB} - H \tag{8}$$

The combined physics-informed loss function is synthesized from Eq. 3,5,6,7,and 8, to impose multiple constraints on the neural network:

$$\mathcal{L} = \sum_{i = \{PDE, BC, IC, EC, OB\}} \lambda_i \left(\frac{1}{N} \sum_{j=1}^N [RES_j]^2 \right)_i$$
(9)

179 Where *j* represents each input data point. By assigning weights λ_i to each residual term, the contri-180 bution of different residual constraints to the total loss function can be adjusted to effectively inform 181 the training process of the neural network.

1822.3 CONSTRAINTS FROM BOUNDARY AND INITIAL CONDITION (HARD CONSTRAINTS)

The optimization process of the neural network relies on the physics-informed loss function, which 184 is centered around the groundwater seepage PDE, as discussed in Section 2.2. The physics-informed 185 loss function involves the interaction between multiple loss function terms, and if multiple losses are solved iteratively without additional processing, the solution mode is referred to as a "soft con-187 straint". In soft constraint, the direction of the neural network solution may deviate due to different 188 solution gradients of tasks, particularly when the boundaries of the objective function or initial con-189 ditions are unconventional. Thus, assigning weights to the loss function for each task is a relative 190 compromise; however, these loss weights are usually manually adjusted in a tedious trial-and-error 191 process which also increases the computational cost. To address this issue, we introduce the "hard 192 constraints" by constructing the neural network output in the form of an elementary function that satisfies Dirichlet boundary conditions, along with the boundary condition equation. Thus, avoiding 193 directly constructing the residual terms RES_{BC} and RES_{IC} for the boundary and initial condi-194 tions, significantly reducing the number of residuals, alleviating the pressure on the model when 195 faced with multiple optimization problems, and facilitating the manual adjustment of training hy-196 perparameters λ_i . Simultaneously, since the hydraulic head processed with hard constraints satisfies 197 the constraints of boundary and initial conditions, the model does not need to find a specific solution under uncertain boundary conditions during the optimization stage, thus enhancing the overall con-199 vergence effect and eliminating the need for labeled data to provide specific solution information. 200

$$\hat{H}_{C} = C(x, y, t)\hat{H} + \varphi(x, y, t), \begin{cases} \varphi(x, y, 0) = H_{0}(x, y) \\ \varphi(x, y, t) = H(x, y, t) \big|_{\Gamma_{1}}, \quad (x, y, t) \in \Gamma_{1} \end{cases}$$
(10)

Where \hat{H}_C is the predicted hydraulic head of the model considering hard constraints; C(x, y, t) is a boundary constraint function characterized by its output gradually tending toward 0 as the input spatial-temporal coordinates (x, y, t) approach the model boundary or initial time. For groundwater dynamics in the river-canal system addressed in this study, since Dirichlet boundary conditions exist only on the left-hand and right-hand sides, the following constraint function can be constructed:

$$C(x, y, t) = \frac{(x - x_{\min})(x_{\max} - x)t}{N_s}$$
(11)

Where x_{min} and x_{max} are the minimum and maximum values of the spatial coordinates in the xaxis direction of the canal (the x-coordinate values of the left and right boundary conditions); N_s is a zoom factor used to balance the data magnitude, which is set as the maximum value of l^2t . With hard constraints applied, the model's output at boundary and initial points directly matches $\varphi(X)$, while other points are constrained by the PDE. Since $\varphi(X)$ inherently satisfies boundary and initial conditions, separate loss functions for these conditions are unnecessary, reducing the overall loss functions and improving the model's solution accuracy.

216 2.4 INPUT FEATURE FUSION

243 244

249

250

In the river-canal groundwater seepage field, the water levels (H_a, H_b) of the left and right canals play a crucial role in the morphology of the phreatic surface, and these water levels are the key boundary conditions for defining specific solutions of the PDE of groundwater seepage. In the baseline PINNs (as discussed in Section 2.2), the convergence of the model depends on fixed boundary conditions to find a specific solution, making it difficult to adjust once the boundary conditions are set. In this study, we proposed to adopt an innovative constraint to the equation which integrates the left and right canal water levels (H_a, H_b) as dynamic input features into the model and utilizes them in the constraint equation, thus limiting the model outputs to a preset boundary condition:

$$\hat{H}_{C} = C(x, y, t)\hat{H}(x, y, t, H_{a}, H_{b}) + \varphi_{BC}(x, y, t, H_{a}, H_{b}), \begin{cases} \varphi_{BC}(x = x_{\min}, y, t) = H_{a} \\ \varphi_{BC}(x = x_{\max}, y, t) = H_{b} \end{cases}$$
(12)

Where φ_{BC} is a condition function of definite solution for input features (discussed in Section 2.3).

229 In the groundwater seepage equation of a river-canal system(Eq. 2), in addition to the spatial-230 temporal coordinate variables (x, y, t), the hydraulic conductivity (K) and the source/sink term (W) 231 are also important parameters affecting groundwater seepage behavior. Incorporating these param-232 eters into the model's input features can provide complete hydrogeological background information 233 for the groundwater seepage equation, thereby improving the model's flexibility and adaptability 234 in simulating different groundwater seepage scenarios. In this study, randomly generated hydraulic 235 conductivities and source/sink terms are added to the training dataset of spatial-temporal coordi-236 nate points, creating a dataset with more features to integrate hydraulic conductivity and source/sink term information into the training process. These additional variables are then introduced into the 237 physics-informed residual term of PDE (Eq. 13), allowing the model to receive and process diverse 238 seepage scenarios described by the hydrogeological parameters. By constructing random datasets, 239 the model can be trained on a large number of generated training cases, each representing a specific 240 hydrogeological condition, thus enhancing the model's ability to simulate different groundwater 241 seepage conditions. 242

$$RES_{PDE}(x, y, t, W, K) = \frac{\partial}{\partial x} \left(K\hat{H}_C \frac{\partial \hat{H}_C}{\partial x} \right) + \frac{\partial}{\partial y} \left(K\hat{H}_C \frac{\partial \hat{H}_C}{\partial y} \right) + W - \mu \frac{\partial \hat{H}_C}{\partial t}$$
(13)

This model can simulate more complex and diverse hydrogeological conditions by incorporating the
hydraulic conductivity, source/sink term, and boundary conditions (left and right canal water levels)
into the model input features, covering a wider range of practical applications, greatly enhancing the
usefulness and flexibility of the model.

3 EXPERIMENT

251 The PI-RGSM model proposed in this study consists of fully connected neural networks (FCNN), 252 hard constraints, and loss function design. As shown in Figure 2 (ignoring the red dashed line in the 253 input feature section), the input features of the model include the spatial-temporal coordinates of the river-canals seepage (x, y, t), adjustable water levels of the river-canals (H_a, H_b) , and hydrogeolog-254 ical parameters (e.g., the source/sink term W and hydraulic conductivity K). These multidimen-255 sional input data are processed through the neural network to output the initial predicted hydraulic 256 head H. The model introduces boundary condition function C(X) and condition function of definite 257 solution $\varphi(X)$ as hard constraints to constrain the model to predict the hydraulic head (H_C) . Au-258 tomatic differentiation (AD) is then employed to compute the partial derivatives of hydraulic head. 259 The source/sink term (W) and hydrogeological parameters (e.g., hydraulic conductivity) are incor-260 porated into the PDE constraint ($LOSS_{PHY}$), along with the engineering control loss ($LOSS_{EC}$), 261 to form a system of PDEs (Ketkar et al., 2021). These two loss terms are balanced to optimize 262 the overall training by adjusting the parameters λ_1 and λ_2 . The model demonstrates flexibility by 263 outputting the corresponding hydraulic head for different boundary conditions and hydrogeological 264 parameters without retraining.

To verify the performance of PI-RGSM, a case involving groundwater seepage in a river-canal system under a unconfined aquifer was established. Without loss generality, the unit of length is set to meter (m), and the unit of time is set to day (d). The schematic of this case is illustrated in Figure 1, where the porous media region between the left and right canals represents the infiltration area. The infiltration area has a length (x) of 40m and a width (y) of 10m. The impermeable bottom boundary is considered to be flat, with a formation thickness (M) is 5m. The specific storage (μ) is 0.1. The



Figure 2: The structure of PI-RGSM and PI-RGSM-K: where the blue line section represents the PI-RGSM, and the red dashed section highlights the extended model PI-RGSM-K, which builds on PI-RGSM by parameterizing K as a spatial function of x and y. 292

293 hydraulic conductivity (K) ranges from 0.5m/d to 0.9m/d. The source/sink term (W) is positive (representing recharge) and varies from 0.000m/d to 0.007m/d. The initial water levels (H_a, H_b) of the left and right canals vary between 2m and 3.5m. 295

The core architecture of PI-RGSM is a deep neural network with 7 layers, each containing 50 neurons, using the Adam optimizer. This configuration was chosen to align with the optimal architecture identified during the hyperparameter tuning process for the baseline PINNs (as detailed in the appendix F). According to the experimental settings of hydrogeological parameters, boundary and initial conditions (hard constraints), and engineering control, the physics-informed losses ($LOSS_{PHY}$ and $LOSS_{EC}$) are defined as follows:

$$\begin{cases} LOSS_{PHY} = RES_{PDE} = \frac{\partial}{\partial x} \left(K\hat{H}_C \frac{\partial \hat{H}_C}{\partial x} \right) + \frac{\partial}{\partial y} \left(K\hat{H}_C \frac{\partial \hat{H}_C}{\partial y} \right) + W - 0.1 \frac{\partial \hat{H}_C}{\partial t} \\ LOSS_{EC} = RES_{EC} = \begin{cases} 2 - \hat{H}_C, & \hat{H}_C < 2 \\ 0, & 2 \le \hat{H}_C \end{cases} \end{cases}$$
(14)

305 306

307

308

309

310

311

312

313

291

296

297

298

299

300

301 302 303

> Since the PI-RGSM model adopts a self-supervised learning approach which does not rely on labeled training data. It requires the generation of random values within specified ranges for each input feature to simulate diverse scenarios due to the scarcity of real hydrogeological data in groundwater systems. Table 1 lists all input features and their corresponding value ranges. These datas were validated through their application in professional simulation software like MODFLOW, ensuring their relevance to real-world conditions. In this experiment, 10,000 training samples were generated, and all input data were normalized. The model was trained using these data for 1,000 iterations.

Table 1: Input features and value ranges of the PI-RGSM

Input features	x(m)	y(m)	t(d)	W(m/d)	K(m/d)	$H_a(m)$	$H_b(m)$
Value ranges	[0,40]	[0,10]	[0,100]	[0,0.007]	[0.5,0.9]	[2,3.5]	[2,3.5]

Although the PI-RGSM model demonstrates potential in predicting groundwater seepage in river-318 canal system, its application is limited by relatively simplified geological environment assumptions, 319 such as the homogeneous hydraulic conductivity which ignores the heterogeneity of the hydraulic 320 conductivity in practical application scenarios. To address complex geological conditions, this study 321 proposes an extension of the PI-RGSM framework (hereafter refer to PI-RGSM-K). 322

To consider the heterogeneity of the hydraulic conductivity field, we replace the homogeneous hy-323 draulic conductivity (a single value for the whole area) as input by regarding the hydraulic conduc324 tivity as a function of the two-dimensional space coordinates x and y (i.e., K(x, y)). This approach 325 allows the model to directly map the hydraulic conductivities and their gradients along the coordi-326 nate axes into a physics-informed loss function, thus incorporating geological features in training 327 process. Based on this improvement, the core structure of the PI-RGSM-K can be directly extended 328 from the PI-RGSM framework (shown in Figure 2). The key modification involves constructing the hydraulic conductivity K as an independent function K(x, y) dependent on two-dimensional spatial 329 coordinates. Along with the source/sink term W in the input features, it is incorporated into the 330 PDE constraint ($LOSS_{PHY}$) to improve prediction under heterogeneous porous media. Compared 331 to PI-RGSM, PI-RGSM-K uses a function K=-0.01x+0.8 as input (data details in appendix E) in 332 our study. However, K can also be parameterized using other functions of x and y, or even as a 333 randomly generated hydraulic conductivity field. When K is a random permeability field that can 334 better reflect the real geological environment, we can build a separate neural network to take (x, y)335 as the input of the neural network, and K as the output. Using the powerful function approximation 336 ability and automatic differential characteristics of the neural network, we can obtain the gradient 337 characteristics of K about x and y, which is the direction we need to improve in the future. 338

4 **RESULTS AND DISCUSSION**

340 4.1 PI-RGSM

339

361

362

364 365 366

341 The major advantage of PI-RGSM is the self-supervised learning scheme. To verify whether the hard 342 constraint method enables PI-RGSM to achieve self-supervised learning and reduce the dependence 343 on observations, experiments were conducted between PI-RGSM and baseline PINNs and their 344 extended models. The experiment did not use observations but instead relied solely on the physics-345 informed loss function for model training. Since the baseline PINNs and their extended models 346 cannot modify the predefined hydrogeological parameters after training, the experiment was set 347 to fixed conditions: precipitation W=0.004m/d, homogeneous hydraulic conductivity K=0.7m/d, 348 and canal heights of 2m. The prediction results were compared with reference data generated by 349 MODFLOW. For the baseline PINNs, the loss term weights were set as $\lambda_{PDE} = 100, \lambda_{BC} = 1$, 350 $\lambda_{IC} = 1$, $\lambda_{EC} = 1$, and $\lambda_{OB} = 1$ based on empirical experience to balance their contributions to the overall loss function. The network structure of baseline PINNs was determined through hy-351 perparameter tuning, with 7 layers and 50 neurons per layer identified as the optimal configuration. 352 Detailed tuning processes and results are provided in the appendix F. 353

354 The performance metrics of models are calculated and shown in Table 2. Results show that the 355 baseline PINNs fail to capture the variation trend of phreatic head, with high errors. Extensions 356 like RCNN and GW-PINN demonstrate significantly better performance compared to the baseline, achieving R^2 values of 0.95862 and 0.96246, respectively, but they are limited to a single hydrogeo-357 logical environment. However, PI-RGSM further outperforms them under the same test conditions, 358 achieving an R^2 accuracy exceeding 0.98. This result demonstrates that the PI-RGSM successfully 359 eliminates dependence on observations and realizes self-supervised learning. 360

Table 2: Performance of baseline PINNs	and PI-RG	SM withou	t observation
	MAE	RMSE	R^2
Baseline PINNs (Raissi et al., 2019)	0.33046	0.36150	-4.32762
RCNN (Sun et al., 2023)	0.13026	0.11265	0.95862
GW-PINN (Zhang et al., 2022)	0.12658	0.09523	0.96246
PI-RGSM	0.01391	0.01651	0.98889

367 Another advantage of the PI-RGSM is its adaptivity to different source/sink terms and hydroge-368 ological parameters for inference without retraining. To verify whether PI-RGSM can accurately 369 capture phreatic surface based on source/sink term and hydrogeological parameter without retraining, the trained PI-RGSM is applied to different source/sink terms and hydrogeological parameters 370 as inputs. The results are compared with reference data generated by MODFLOW. 371

372 (1) The influence of precipitation and hydraulic conductivity: To evaluate the performance of the 373 PI-RGSM under different source/sink terms and hydrogeological conditions, a series of experiments 374 was designed to simulate different groundwater seepage scenarios by adjusting precipitation W and 375 hydraulic conductivity K. All experiments were carried out under fixed left and right canal water levels $H_a=H_b=2m$ to ensure consistent boundary conditions. Various combinations of hydrogeo-376 logical parameters were set up in the experiment. First, the effect of source/sink terms (precipitation 377 W) on the model was tested under a fixed hydraulic conductivity of K=0.700m/d with three differ-

ent precipitation levels: low, medium, and high (W=0.002m/d, W=0.004m/d), and W=0.006m/d). Then, the hydraulic conductivity was adjusted, and three different values were tested (K=0.500m/d), K=0.700m/d, and K=0.900m/d) under a constant precipitation rate of W=0.005m/d.

The test results are shown in Table 3. For each set of test configurations, the hydraulic head predicted by the model is displayed as a 3D graph (Figure 3 and 4), with the phreatic line profile extracted at y=5m for comparison with the reference data generated by MODFLOW. The 3D image on the top displays the prediction of the phreatic surface, while the profile on the bottom shows the phreatic line at y=5m; The blue curve represents the model prediction, and the red curve denotes the refer-ence curve from MODFLOW. The highest hydraulic head obtained from the prediction curve and the reference curve are marked in the figure. The results demonstrate that the phreatic surface pre-dicted by the model is smooth, consistent with the typical morphology of groundwater seepage in river-canal system, and almost completely matches the reference curve from MODFLOW. In Figure 3, as precipitation increases from 0.002m/d to 0.006m/d, the model predicts a corresponding rise in hydraulic head. For example, at the lowest precipitation (0.002m/d), the predicted maximum hydraulic head is approximately 2.29m, while at the highest precipitation (0.006m/d), it rises to about 2.75m, demonstrating the model's sensitivity and adaptability to changes in precipitation. In the case of different hydraulic conductivity (Figure 4), PI-RGSM also shows a high degree of accu-racy, with a mean R^2 exceeding 0.977 relative to the reference data. These experiments confirm the high adaptability and accuracy of the model under different source/sink terms and hydrogeological conditions and its physical consistency and reliability in capturing groundwater seepage behavior. K = 0.700(m/d), W = 0.006(m/d)K = 0.700(m/d), W = 0.002(m/d)K = 0.700 (m/d), W = 0.004 (m/d)







430 (2) The influence of boundary condition changes: A series of experiments were designed to 431 evaluate the response of the PI-RGSM to changes in boundary conditions, maintaining precipitation W and hydraulic conductivity K at 0.004m/d and 0.700m/d, respectively, to ensure consistency in hydrogeological parameters. The key variables in this experiment are the water levels of the left and right canals (i.e., H_a and H_b). Here, H_b is kept constant at 2.000*m*, while H_a varied at 2.500*m*, 3.000*m*, and 3.500*m*. This configuration aims to simulate diverse water level differences to examine the model performance under varying boundary conditions.

The test results are summarized in Table 4. To visualize the prediction performance of the model under different boundary water levels, 3D graphs are plotted. As H_a gradually increases from 2.500m to 3.500m, the predicted hydraulic head changes accordingly (shown in Figure 5). In particular, the profile comparison at y=5m reveals that the predicted hydraulic head gradually increases with the elevation of H_a , maintaining a high degree of consistency with the MODFLOW reference. These results demonstrate the response of the PI-RGSM under changing boundary conditions and emphasizes its high prediction accuracy, with an mean R^2 exceeding 0.99.



Figure 5: Phreatic surfaces from PI-RGSM under different boundary conditions.

///= "						
401	Input f	Input features		RMSE	R^2	
462	$H_a = H_b$	= 2.000				
463		W = 0.002	0.01621	0.01847	0.95065	
464	K = 0.700	W = 0.004	0.01391	0.01651	0.98889	
465		W = 0.006	0.02639	0.03064	0.98110	
466		K = 0.500	0.03388	0.03934	0.97598	
467	W = 0.003	K = 0.700	0.02013	0.02394	0.98420	
468		K = 0.900	0.01685	0.01938	0.98390	
469	Table 4: Performan	ce of PI-RGS	M under dif	ferent bour	dary condit	ions
470						
471	Input f	Input features		RMSE	R^2	
170	W = 0.004	K = 0.700				
4/2	$H_a = 2.500.$	$H_b = 2.000$	0.02189	0.02392	0.99410	
473	$H_{a} = 3.000$	$H_{b} = 2.000$	0.01701	0.02077	0.98926	
474	$H_a = 3.500$	$H_{\rm h} = 2.000$	0.04012	0.04411	0.99002	
475						

Table 3: Performance of PI-RGSM under different source/sink terms and hydrogeological

476 4.2 PI-RGSM-K

443

450

451

452

453

454

455 456

457

458

459 460

To evaluate the performance of the PI-RGSM-K in handling heterogeneous hydraulic conductivity
fields, experiments was conducted by under different conditions of source/sink terms and water
levels of left and right canals. Under the complex geological conditions of heterogeneous hydraulic
conductivity field, the performance metrics were calculated and shown in Table 5. The predicted
phreatic surface was visualized using 3D graphs and cross-sectional diagrams of the phreatic line
(Figure 6 and 8). Vertical and dashed lines indicated the positions of the predicted and referenced
maximum hydraulic heads.

In the experiment different precipitation (Figure 6), the phreatic surface predicted by PI-RGSM-K rose accordingly as precipitation increased from 0.003m/d to 0.007m/d, especially under high precipitation conditions. This demonstrates that PI-RGSM-K can accurately capture the impact of

486 precipitation change on the phreatic surface in heterogeneous hydraulic conductivity field. The pre-487 dicted results are highly consistent with the reference curve of MODFLOW, with R^2 exceeding 488 0.99. In the experiment in response to boundary condition changes (results in appendix G), the hy-489 draulic head predicted by the model changes correspondingly as the water level H_a increases from 490 2.500m to 3.500m. In the profile comparison at y=5m, it is evident that the predicted hydraulic head increases with the elevation of H_a . The results indicate a high degree of consistency with the 491 MODFLOW reference, with a mean $\overline{R^2}$ above 0.98. It is further proven that the PI-RGSM-K has 492 strong adaptability and accuracy in processing boundary conditions, especially under heterogeneous 493 hydraulic conductivity field, and can accurately capture changes in phreatic surface morphology 494 caused by river-canal water levels changes. Additionally, when responding to changes in precipita-495 tion and boundary conditions, it can be observed that the groundwater flow increases on the left side 496 as the hydraulic conductivity field gradually decreases from the left side to the right side. This result 497 in the shift of the peak value of the phreatic surface to the right side, as clearly seen in the figures. 498 This high level of physical consistency proves its superior performance in groundwater seepage pre-499 diction tasks involving heterogeneous hydraulic conductivity field. The performance remains stable 500 under various hydrogeological conditions, which demonstrate the potential application of the model in future groundwater resource management and hydrology research. Future studies will further 501 explore the applicability and optimization methods of the model in a wider range of groundwater 502 seepage scenarios.



Table 5: Performance of PI-RGSM under different precipitations and boundary conditions



504 505

506

507

509

510

511

512 513 514

515

516 517 518

519

521

522

523

524 525

527

528 529

530 This study proposes the PI-RGSM and its extension PI-RGSM-K based on PINNs which signifi-531 cantly improve the accuracy and adaptability of groundwater seepage prediction. The main con-532 tributions are as follows: (1) Self-supervised learning of PI-RGSM is achieved by integrating hard constraints of boundary and initial conditions, reducing the dependence on observations. (2) PI-534 RGSM demonstrates strong generalization capabilities, as it can adapt to varying hydrogeological conditions without requiring retraining, significantly improving its practicality in real-world ap-536 plications. (3) PI-RGSM-K is proposed to simulate heterogeneous porous media by considering 537 heterogeneous hydraulic conductivity. This study demonstrates that integrating physical knowledge into a neural network framework provides an effective scientific method for the sustainable manage-538 ment and protection of groundwater resources, highlighting the great potential and practical value of the physics-informed method in groundwater seepage modeling.

540 REFERENCES 541

559

561 562

563

564

565

566

567

581

- Zoran Babovi and Branislav Bajat. Research in computing-intensive simulations for nature-oriented 542 civil-engineering and related scientific fields, using machine learning and big data: an overview 543 of open problems. Journal of Big Data, 10(1):73, 2023. 544
- Philip Brunner, Craig T Simmons, Peter G Cook, and René Therrien. Modeling surface water-546 groundwater interaction with modflow: Some considerations. Groundwater, 48(2):174-180, 547 2010. 548
- Silvina Caíno-Lores, Andrei Lapin, Jesús Carretero, and Petter Kropf. Applying big data paradigms 549 to a large scale scientific workflow: Lessons learned and future directions. Future Generation 550 Computer Systems, 110:440-452, 2020. 551
- 552 Chong Chen, Wei He, Han Zhou, Yaru Xue, and Mingda Zhu. A comparative study among machine 553 learning and numerical models for simulating groundwater dynamics in the heihe river basin, 554 northwestern china. Scientific reports, 10(1):3904, 2020. 555
- Kyeungwoo Cho and Yeonjoo Kim. Improving streamflow prediction in the wrf-hydro model with 556 lstm networks. Journal of Hydrology, 605:127297, 2022.
- 558 Anshul Choudhary, John F Lindner, Elliott G Holliday, Scott T Miller, Sudeshna Sinha, and William L Ditto. Physics-enhanced neural networks learn order and chaos. Physical Review 560 *E*, 101(6):062207, 2020.
 - George Cybenko. Approximation by superpositions of a sigmoidal function. Mathematics of control, signals and systems, 2(4):303-314, 1989.
 - Li Daolun, Shen Luhang, Zha Wenshu, Liu Xuliang, and Tan Jieqing. Physics-constrained deep learning for solving seepage equation. Journal of Petroleum Science and Engineering, 206: 109046, 2021.
- 568 AD Galeev, EV Starovoytova, and SI Ponikarov. Numerical simulation of the consequences of 569 liquefied ammonia instantaneous release using fluent software. Process safety and environmental 570 protection, 91(3):191–201, 2013.
- 571 Ehsan Haghighat, Maziar Raissi, Adrian Moure, Hector Gomez, and Ruben Juanes. A physics-572 informed deep learning framework for inversion and surrogate modeling in solid mechanics. 573 Computer Methods in Applied Mechanics and Engineering, 379:113741, 2021. 574
- 575 Arlen W Harbaugh. MODFLOW-2005, the US Geological Survey modular ground-water model: 576 the ground-water flow process, volume 6. US Department of the Interior, US Geological Survey 577 Reston, VA, USA, 2005.
- 578 V Hariharan and M Uma Shankar. A review of visual modflow applications in groundwater mod-579 elling. In IOP Conference Series: Materials Science and Engineering, volume 263, pp. 032025. 580 IOP Publishing, 2017.
- 582 Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural 583 networks. science, 313(5786):504-507, 2006. 584
- Hao Jing, Xin He, Yong Tian, Michele Lancia, Guoliang Cao, Alessandro Crivellari, Zhilin Guo, 585 and Chunmiao Zheng. Comparison and interpretation of data-driven models for simulating site-586 specific human-impacted groundwater dynamics in the north china plain. Journal of Hydrology, 587 616:128751, 2023. 588
- 589 Nikhil Ketkar, Jojo Moolayil, Nikhil Ketkar, and Jojo Moolayil. Automatic differentiation in deep 590 learning. Deep Learning with Python: Learn Best Practices of Deep Learning Models with Py-591 Torch, pp. 133–145, 2021.
- Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. Communications of the ACM, 60(6):84–90, 2017.

605

631

- Jie Li, Wenke Wang, Dawei Cheng, Ying Li, Ping Wu, and Xiaoqin Huang. Hydrogeological structure modelling based on an integrated approach using multi-source data. *Journal of Hydrology*, 600:126435, 2021.
- Xin Li, Min Feng, Youhua Ran, Yang Su, Feng Liu, Chunlin Huang, Huanfeng Shen, Qing Xiao,
 Jianbin Su, Shiwei Yuan, et al. Big data in earth system science and progress towards a digital
 twin. *Nature Reviews Earth & Environment*, 4(5):319–332, 2023.
- Sylvain Massuel, Jeanne Riaux, François Molle, Marcel Kuper, Andrew Ogilvie, A-L Collard, Christian Leduc, and Olivier Barreteau. Inspiring a broader socio-hydrological negotiation approach with interdisciplinary field-based experience. *Water Resources Research*, 54(4):2510– 2522, 2018.
- Reed M Maxwell, Laura E Condon, and Peter Melchior. A physics-informed, machine learning
 emulator of a 2d surface water model: what temporal networks and simulation-based inference
 can help us learn about hydrologic processes. *Water*, 13(24):3633, 2021.
- Rakesh Kumar Mishra. Fresh water availability and its global challenge. *British Journal of Multidisciplinary and Advanced Studies*, 4(3):1–78, 2023.
- Sarvin Moradi, Burak Duran, Saeed Eftekhar Azam, and Massood Mofid. Novel physics-informed
 artificial neural network architectures for system and input identification of structural dynamics
 Buildings, 13(3):650, 2023.
- 615
 616
 617
 COMSOL Multiphysics. Introduction to comsol multiphysics[®]. COMSOL Multiphysics, Burlington, MA, accessed Feb, 9(2018):32, 1998.
- 618
 619
 620
 CC Niu, SB Li, JJ Hu, YB Dad, Z Cao, and X Li. Application of machine learning in material informatics: A survey. *Materials Reports*, 34(23):23100–23108, 2020.
- Amin Pashaei Kalajahi, Isaac Perez-Raya, and Roshan M D'Souza. Physics informed deep neural
 net inverse modeling for estimating model parameters in permeable porous media flows. *Journal of Fluids Engineering*, 144(6):061102, 2022.
- Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.
- 628 Chao Song, Tariq Alkhalifah, and Umair Bin Waheed. Solving the frequency-domain acoustic vti
 629 wave equation using physics-informed neural networks. *Geophysical Journal International*, 225
 630 (2):846–859, 2021.
- Guangli Su, Chunbao Xiong, Guiying Zhang, Yong Wang, Qiang Shen, Xiaohu Chen, Hewen An, and Longchao Qin. Coupled processes of groundwater dynamics and land subsidence in the context of active human intervention, a case in tianjin, china. *Science of the Total Environment*, 903:166803, 2023.
- Jian Sun, Xungui Li, Qiyong Yang, Yi Tian, Shaobo Wang, and Meiqing Yang. Hydrodynamic numerical simulations based on residual cooperative neural network. *Advances in Water Resources*, 180:104523, 2023.
- ⁶³⁹
 ⁶⁴⁰ Zhong Yi Wan, Pantelis Vlachas, Petros Koumoutsakos, and Themistoklis Sapsis. Data-assisted reduced-order modeling of extreme events in complex dynamical systems. *PloS one*, 13(5): e0197704, 2018.
- Jiaji Wang, YL Mo, Bassam Izzuddin, and Chul-Woo Kim. Exact dirichlet boundary physics informed neural network epinn for solid mechanics. *Computer Methods in Applied Mechanics* and Engineering, 414:116184, 2023.
- 647 Nanzhe Wang, Dongxiao Zhang, Haibin Chang, and Heng Li. Deep learning of subsurface flow via theory-guided neural network. *Journal of Hydrology*, 584:124700, 2020.

- Nanzhe Wang, Haibin Chang, and Dongxiao Zhang. Deep-learning-based inverse modeling approaches: A subsurface flow example. *Journal of Geophysical Research: Solid Earth*, 126(2): e2020JB020549, 2021a.
- Nanzhe Wang, Haibin Chang, and Dongxiao Zhang. Theory-guided auto-encoder for surrogate
 construction and inverse modeling. *Computer Methods in Applied Mechanics and Engineering*, 385:114037, 2021b.
 - Xianjun Xie, Jianbo Shi, Kunfu Pi, Yamin Deng, Bing Yan, Lei Tong, Linlin Yao, Yiran Dong, Junxia Li, Liyuan Ma, et al. Groundwater quality and public health. *Annual Review of Environment and Resources*, 48(1):395–418, 2023.
 - Kai Yu, Lei Jia, Yuqiang Chen, Wei Xu, et al. Deep learning: yesterday, today, and tomorrow. *Journal of computer Research and Development*, 50(9):1799–1804, 2013.
 - Xiaoping Zhang, Yan Zhu, Jing Wang, Lili Ju, Yingzhi Qian, Ming Ye, and Jinzhong Yang. Gwpinn: A deep learning algorithm for solving groundwater flow equations. *Advances in Water Resources*, 165:104243, 2022.

A RELATED WORKS

651

655

656

657

658

659

660 661

662

663

664 665 666

667

668 With the evolution of scientific paradigms (Caíno-Lores et al., 2020), researches on natural science 669 has mitigated from empirical and theoretical approaches to computational and data-intensive sci-670 entific paradigms driven by computer simulations and big data (Babovi & Bajat, 2023; Li et al., 671 2023). In the field of computational fluid dynamics, numerical models including FLUENT (Galeev et al., 2013), MODFLOW (Brunner et al., 2010; Harbaugh, 2005; Hariharan & Shankar, 2017), and 672 COMSOL Multiphysics (Multiphysics, 1998), have achieved accurate simulations of physical pro-673 cesses like complex multiphase flow, solute transport, and interactions between surface and soil wa-674 ter within groundwater systems. Important issues in those spatial discretization methods for solving 675 the partial differential equations (PDEs) of groundwater are the computationally intensive. More-676 over, the implementation of these models is typically challenging for personnel at water resource 677 management agencies, as highly specialized knowledge and extensive experience in numerical mod-678 eling are required. Continuous advancements in data-intensive scientific paradigms have empowered 679 deep neural networks to achieve significant breakthroughs in image recognition (Krizhevsky et al., 680 2017), meteorological modeling (Wan et al., 2018), material informatics (Niu et al., 2020), hydrol-681 ogy (Maxwell et al., 2021), and mathematics(Choudhary et al., 2020), due to the powerful feature 682 extraction (Yu et al., 2013; Hinton & Salakhutdinov, 2006) and complex function approximation abilities (Cybenko, 1989). Nonetheless, as "black box" models, neural networks exhibit a lack of 683 transparency in their decision-making processes and depend significantly on extensive training data, 684 limiting their use in groundwater research. Recently, by combining physical laws with neural net-685 works, the physics-informed neural network (PINNs) method has improved the interpretability of 686 models and diminished reliance on extensive datasets, providing an efficient and accurate approach 687 for groundwater system modeling, and thus promoting the scientific and reliability of groundwater 688 prediction and management decisions. 689

The concept of PINNs was first proposed by Raissi et al. (2019). In their approach, the neural net-690 work was trained by constructing physical constraints, and optimized by using the residuals based 691 on initial conditions, boundary conditions, and governing equations as loss functions. Neural net-692 work derivatives were computed through automatic differentiation (AD). This method effectively 693 solves both continuous and discrete PDEs, marking a new era of integration of physical knowl-694 edge and deep learning. The introduction of PINNs has significantly expanded the scope of neural networks in scientific research and engineering applications, while markedly enhancing the predic-696 tion accuracy, generalization ability, and interpretability of models. Since then, researchers have 697 continued to explore and innovate within the PINNs framework, applying it to various fields such 698 as fluid mechanics (Haghighat et al., 2021), structural dynamics (Moradi et al., 2023), acoustics (Song et al., 2021), and solid mechanics (Wang et al., 2023). In hydrological research, PINNs of-699 fer a new perspective for more convenient and efficient prediction of groundwater flow with their 700 unique physical guidance characteristics. Wang et al. (2020) proposed a theoretically guided neural 701 network model (TgNN) based on the loss function of groundwater seepage differential equations,

702 which improved the generalization performance of the groundwater seepage model with limited 703 observational data. Additionally, the team demonstrated that the physics-informed model could ef-704 fectively use physical information to predict groundwater seepage responses beyond the training set 705 through groundwater seepage transfer learning tasks. Cho & Kim (2022) used an LSTM network 706 to regress the residual differences between a neural network model and the Weather Research and Forecasting model-Hydro (WRF-Hydro). They developed a hybrid model (WRF-Hydro-LSTM) to predict groundwater levels, and experimental results indicated that this hybrid model achieved higher 708 prediction accuracy, less sensitivity to training datasets, and better generalization performance than 709 the LSTM network or WRF-Hydro model alone. Wang et al. (2021a) proposed a neural network 710 constrained by geostatistical information and a physics-guided autoencoder based on convolutional 711 neural networks (CNNs) (Wang et al., 2021b), applying the physical guidance method to the pa-712 rameter inversion of groundwater hydraulic conductivity fields. The model's accuracy and practi-713 cality were further enhanced. The study by (Pashaei Kalajahi et al., 2022) demonstrated that, even 714 with limited data, physics-guided neural networks could accurately estimate key parameters such 715 as hydraulic conductivity and porosity of the medium, showcasing their strong ability to combine 716 data-driven approaches with physical guidance. Daolun et al. (2021) established a new specialized 717 neuron model incorporating pressure gradient information and proposed an algorithm called signpost neural network (SNN), which significantly improves the accuracy of solving unsteady seepage 718 partial differential equations. Although significant progress has been made in improving groundwa-719 ter seepage models using PINNs, current models still heavily depend on the adequacy and quality 720 of observed data and remain sensitive to outliers. If the training datasets are insufficient or contain 721 corrupted data, the simulation results may exhibit considerable bias. Moreover, after the model has 722 been trained, its applicability is generally limited to specific hydrogeological parameter settings, 723 making it difficult to generalize to broader or unforeseen hydrogeological scenarios, thus restricting 724 the model's versatility and practicality. 725

726 727

728

729

730 731

732

739

740 741

749

753

B BASIC LAW OF GROUNDWATER SEEPAGE

Without considering changes in water density, the flow of groundwater in a three-dimensional aquifer within porous media can be expressed by the following partial differential equation:

$$\frac{\partial}{\partial x}\left(K_{xx}\frac{\partial H}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_{yy}\frac{\partial H}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_{zz}\frac{\partial H}{\partial z}\right) + W = \mu_s M \frac{\partial H}{\partial t}$$
(15)

733 Where t is time (T); K_{xx} , K_{yy} , and K_{zz} are the hydraulic conductivities (LT^{-1}) along the x, y, 734 and z axes, respectively. H is the hydraulic head (L) at the corresponding space-time point; W is 735 the source/sink term of groundwater (such as precipitation, pumping, etc.); μ_s is the specific storage 736 (L^{-1}) of the aquifer below the free surface. M is the thickness of the aquifer. Combined with fixed 737 solution conditions (boundary conditions and initial conditions), it can reflect the water balance 738 relationship per unit volume and per unit time under Darcy flow conditions.

C DEEP NEURAL NETWORK

The Deep Neural Network (DNN) consists of an input layer, several hidden layers, and an output layer, where each layer contains multiple neurons connected by a weight matrix (which may include bias terms). Consider a deep neural network with L layers, denoted as $Y = f_{net}(X, \theta)$, with n^l neurons in the L-th layer. The number of neurons in the input layer equals the dimension of the input feature vector X, and the number of neurons in the output layer equals the dimensionality of the output vector Y. The value of the *i*-th neuron in L-th layer, $z_{l,i}$, is computed as the product of the weight $W_{l,i}^T$ and neurons from the previous layer, plus the bias term $b_{l,i}$:

$$z_{l,i} = W_{l,i}^T a_{l-1} + b_{l,i} \tag{16}$$

The output of neurons is nonlinearly transformed by the activation function, which provides nonlinear capability to the network. After processing $z_{l,i}$ with the activation function σ , the output of the *i*-th neuron in the L-th layer, $a_{l,i}$, Lis obtained:

$$a_{l,i} = \sigma(z_{l,i}) \tag{17}$$

The final output of the network is produced by the output layer of the multi-layer neural network, and the network training process relies on a loss function that quantitatively assesses the difference between the network's predictions and the actual target values.

D PERFORMANCE METRICS

To evaluate the performance of the model, three main indicators were used for quantitative analysis: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Determination Coefficient (R^2). Their definitions are as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$
(18)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$
(19)

$$R^{2} = 1 - \frac{\sum_{i=1}^{n} (y_{i} - \hat{y}_{i})^{2}}{\sum_{i=1}^{n} (y_{i} - \bar{y})^{2}}$$
(20)

771 Where n represents the total number of samples, y_i represents the reference output from MODFLOW 772 of the *i*-th sample, \hat{y}_i represents the predicted value of the *i*-th sample by the DNN model, and \bar{y}_i 773 is the average of the reference values of all samples. Ideally, the closer MAE and RMSE are to 0, 774 the smaller the prediction error and the higher the accuracy of the model. The closer the R^2 value is 775 to 1, the better prediction the DNN model achieves.

776 777

756

758

759

760

E DETAILS OF PI-RGSM-K MODEL

778
779 Given that the PI-RGSM-K model is a specific application of the PI-RGSM model under conditions
780 of a heterogeneous hydraulic conductivity field, and since the structures of both models are similar,
781 most of the experimental settings are identical to PI-RGSM as follows:

(1) As the hydraulic conductivity field has been integrated into the physics-informed loss function, the model input features are spatial-temporal coordinates (x, y, t), canal water levels (H_a, H_b) , and source/sink term W. The model still uses a random method to generate input data for self-supervised training (shown in Table 6).

(2) A heterogeneous hydraulic conductivity field is introduced, with K defined as a function that varies linearly with x (i.e., K = -0.01x + 0.8). The hydraulic conductivity linearly changes from 0.8m/d at the left boundary to 0.4m/d at the right boundary.

Table 6: Input features and value ranges of the PI-RGSM-K

Input features	x(m)	y(m)	t(d)	W(m/d)	$H_a(m)$	$H_b(m)$
Value ranges	[0,40]	[0,10]	[0,100]	[0,0.007]	[2,3.5]	[2,3.5]

794 795 796

797

798

791 792 793

F HYPERPARAMETER TUNING PROCEDURE OF BASELINE PINNS

799 The hyperparameter settings of the neural network, including the learning rate, layers, and neurons, 800 directly influence the predictive accuracy of baseline PINNs. Figure 7 illustrates the effects of these factors on the L^2 relative error. As the initial learning rate decreases, the error drops significantly 801 and stabilizes below 1%, reaching a minimum at a learning rate of 0.001. However, further reducing 802 the learning rate introduces additional errors. The results show that increasing the number of layers 803 generally improves prediction accuracy, but the improvement becomes negligible beyond 5 layers. 804 Similarly, increasing the number of neurons per layer reduces the error, with the most significant 805 reduction observed up to 30 neurons per layer. After this point, further increases in neurons yield 806 diminishing returns. 807

The optimal network structure identified in this study consists of 7 layers with 50 neurons per layer, achieving an L^2 error of 2.83%. While deeper or wider networks may further improve accuracy slightly, the additional computational cost outweighs the marginal benefits due to the relatively low complexity of the governing equations in this study. In summary, Too many or too few neurons in hidden layers can cause overfitting or underfitting. Excessive neurons may lead to overfitting due to insufficient training data, while also increasing training time without achieving better performance. Selecting an appropriate number of neurons is essential for balancing accuracy and efficiency.



Figure 7: Impact of learning rate (left) and network parameters (Layers and Neurons) on the performance of baseline PINNs.

PHREATIC SURFACES FROM PI-RGSM-K UNDER DIFFERENT BOUNDARY G CONDITIONS.



Figure 8: Phreatic surfaces from PI-RGSM-K under different boundary conditions.