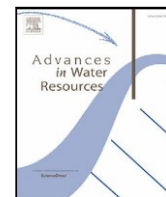




Contents lists available at Science Direct

## Advances in Water Resources

journal homepage: [www.elsevier.com/locate/advwatres](http://www.elsevier.com/locate/advwatres)

## Boundary-to-solution mapping for groundwater flows in a Toth basin

Jingwei Sun <sup>a</sup>, Jun Li <sup>b</sup>, Yonghong Hao <sup>c</sup>, Cuiting Qi <sup>d</sup>, Chunmei Ma <sup>a</sup>, Huazhi Sun <sup>a</sup>, Negash Begashaw <sup>e</sup>, Gurcan Comert <sup>e</sup>, Yi Sun <sup>f</sup>, Qi Wang <sup>f,\*</sup>

<sup>a</sup> School of Computer Science, Tianjin Normal University, Tianjin, 100084, China <sup>b</sup> School of Mathematical Science, Tianjin Normal University, Tianjin, 300387, China <sup>c</sup> Center for Ground Water Research, Tianjin Normal University, Tianjin, 100084, China <sup>d</sup> School of Geographical and Environmental Science, Tianjin Normal University, Tianjin, 100084, China <sup>e</sup> Computer Science, Physics and Engineering Department, Benedict College, Columbia, SC 20204, USA <sup>f</sup> Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA

## ARTICLE

## INFO

## ABSTRACT

## Keywords:

Machine learning  
Poisson equation  
Boundary-to-solution mapping  
Toth basin  
DeepONet  
Ground water flows

In this paper, the authors propose a new approach to solving the groundwater flow equation in the Toth basin of arbitrary top and bottom topographies using deep learning. Instead of using traditional numerical solvers, they use a DeepONet to produce the boundary-to-solution mapping. This mapping takes the geometry of the physical domain along with the boundary conditions as inputs to output the steady state solution of the groundwater flow equation. To implement the DeepONet, the authors approximate the top and bottom boundaries using truncated Fourier series or piecewise linear representations. They present two different implementations of the DeepONet: one where the Toth basin is embedded in a rectangular computational domain, and another where the Toth basin with arbitrary top and bottom boundaries is mapped into a rectangular computational domain via a nonlinear transformation. They implement the DeepONet with respect to the Dirichlet and Robin boundary condition at the top and the Neumann boundary condition at the impervious bottom boundary, respectively. Using this deep-learning enabled tool, the authors investigate the impact of surface topography on the flow pattern by both the top surface and the bottom impervious boundary with arbitrary geometries. They discover that the average slope of the top surface promotes long-distance transport, while the local curvature controls localized circulations. Additionally, they find that the slope of the bottom impervious boundary can seriously impact the long-distance transport of groundwater flows. Overall, this paper presents a new and innovative approach to solving the groundwater flow equation using deep learning, which allows for the investigation of the impact of surface topography on groundwater flow patterns.

## 1. Introduction

The Toth groundwater flow analysis was a seminal theoretical attempt to relate surface topography of the water table and the associated hydrological boundary conditions with the steady state groundwater flow field driven by gravity in a small drainage basin, known as the Toth basin (Tóth, 1962, 1963). It involved solving an elliptic boundary value problem for a given surface topography of the water table not far from a horizontally flat surface with the associated Dirichlet boundary condition on a rectangular domain approximately. For a general nonrectangular drainage basin with a surface topography far from a flat surface, the elliptic boundary value problem would have to be solved numerically. The Toth water table analysis demonstrated the impact of surface topography and the associated water potential at the boundary on the ground water flow in the basin domain approximately.

the solution of groundwater flow equations only produces one solution for any given boundary conditions. When one studies another Toth

Traditionally, a numerical solver (such as MODFLOW, COMSOL, etc.) for basin with a different surface topography, the solution have to be recalculated completely. Given the flow equation and the boundary condition, the mapping from the boundary condition to the solution is essentially provided by the numerical solver. One thus wonders if the numerical solver can be replaced by a concrete function or “mapping” that is fully capable of producing the solution from any prescribed surface topography without being recalculated.

In the past, the Toth theory was refined through adjusting the coefficient of permeability or the viscosity of the fluid in porous media, extended to examine the influence of temperature (An et al., 2014), and used to investigate the influence of depth and systemic heterogeneity in porous media (Cardenas and Jiang, 2010; Jiang et al., 2011). Several studies focused on generalizing the

\* Corresponding author.

E-mail address: [qwang@math.sc.edu](mailto:qwang@math.sc.edu) (Q. Wang).

<https://doi.org/10.1016/j.advwatres.2023.104448>

Received 15 October 2022; Received in revised form 22 March 2023; Accepted 25 April 2023 Available online

4 May 2023

0309-1708/© 2023 Elsevier Ltd. All rights reserved.

Toth theory to other settings, like the more realistic three-dimensional space (Wang et al., 2017), unsteady situations (Niu et al., 2015), etc. However, none of the studies paid attention to the impact of the top surface topography of the water table on the solution in the Toth basin holistically when it is of an arbitrary shape.

In this paper, we extend the Toth water table study to a domain with an arbitrary piecewise smooth top and bottom boundaries and two physically relevant boundary conditions, and propose a novel approach to establish a mapping from surface topographies (top alone or top+ bottom) to the solution of the groundwater flow equation in the Toth basin directly using a deep learning approach. To some extent, this is an analogue of a solution formula for an initial-boundary value problem for partial differential equations (PDEs) in the context of deep learning, where the solution of the initial-boundary value problem is expressed as a neural network function of the domain, the boundary conditions and the nonhomogeneous forcing term. This approach produces a solution mapping that maps the prescribed initial and boundary conditions as well as the forcing term to the solution directly. It can be readily applied to any geophysical basins that share the same hydrological property such as the mobility/conductivity coefficient in the flow equation. As a demonstration of the approach, we present the mapping while neglecting the forcing effect due to the source or sink and assuming the porous media is spatially homogeneous. We remark that the method applies to any inhomogeneous porous media and groundwater flow equations with a source or sink term. The advantage of this approach is that once the solution mapping is obtained in one Toth basin, it can be readily applied to all other Toth basins where the hydrological property of the porous media is the same, but the boundary and boundary conditions can be different.

The recent advancement in deep learning with neural networks makes the development of such a desired mapping plausible (Karniadakis et al., 2021; Lu et al., 2019; Cao, 2021; Guibas et al., 2022; Kissas et al., 2022; Li et al., 2021a; Pang et al., 2019). Given that a neural network is a mapping composed of compound functions with specific layered structures, the mapping can be established should we propose the proper architecture of the deep neural network in principle in the context of physics-informed machine learning (PIML) (Leshno et al., 1993; Chen and Chen, 1995). We note that the steady state groundwater flow equation in porous media is an elliptic (or Poisson) equation. Given the boundary and physically consistent boundary conditions, a solution can be represented by an integral containing the Green's function (Haberman, 2013). The integral with the Green's function yields the mapping from the boundary, boundary conditions and the source term to the solution theoretically. Motivated by this connection between the domain, boundary conditions and the source term of the equation, we represent the mapping using a new form of neural network, known as the DeepONet. The DeepONet has been shown to have the capacity to establish the mapping between the model parameters, its boundaries (including boundary conditions) to the solution in the domain (Lu et al., 2021). It is therefore an appropriate and powerful tool for us to build the desired boundary-to-solution mapping.

Specifically, we will answer the following questions in this study using machine learning with DeepONet.

- What is the influence of the surface topography and the geometry of the Toth basin to the steady flow field through the water potential in Toth basin  $\mathcal{D}$ ?
- What is the specific effect of both the top and bottom boundary conditions to the solution of the groundwater flow equation in the Toth basin through the boundary-to-solution mapping? We will focus on two types of top boundary conditions: (i) the Dirichlet boundary condition in which the water potential is prescribed at the top boundary related to the altitude of the location:  $h = g\phi(\mathbf{x})$ , where  $h$  is the water potential,  $g$  is gravity, and  $\gamma = \phi$  defines the top boundary; and (ii) the Robin boundary condition:

$\frac{\partial h}{\partial n}$

$$-\frac{\partial h}{\partial n}(\mathbf{x}) + \gamma h(\mathbf{x}) = \gamma g\phi(\mathbf{x}), \text{ where } \gamma \text{ is a rate parameter whose}$$

reciprocal

represents the penetration length. The latter simply states a balance law between the cross boundary flux and the difference between the water potential and a saturated water potential at the top surface. As  $\gamma \rightarrow \infty$ , i.e. the penetration length shrinks to zero, the Dirichlet boundary condition is recovered. Thus, the Robin boundary condition is an approximation to the Dirichlet boundary condition at large  $\gamma \gg 1$ .

- What is the surface topography and basin geometry to the steady state Darcy velocity field in the basin? We note that the topography here refers to the water table topography or water head profile not the topography of the ground surface.

We will address these issues holistically by solving the PDE boundary value problem with respect to two distinct boundary conditions using DeepONet (Pang et al., 2019; Lu et al., 2019). The presentation is given for steady states without a source term. However, we emphasize again that the method extends readily to flow phenomena with a source and transient situations with a minimum modification to the DeepONet architecture. We note that for a completely new Toth basin, an analogous boundary-to-solution mapping to describe ground water flows in the porous medium can be obtained through the transfer learning, which could accelerate the training process and be efficiently done.

Numerically, we present three implementations of the DeepONet in which the boundaries of arbitrary shapes are represented using a piecewise linear interpolant, a truncated Fourier series, or mapped to a flat surface via a nonlinear transformation. All three implementations yield comparable numerical results. Without loss of generality, we will detail the latter two implementations in this paper.

## 2. Mathematical formulation

We first present the model derivation and give a brief discussion on consistency of boundary conditions with the governing equation. Then, we discuss how the solution of the boundary value problem of the steady state groundwater flow equation depends on prescribed boundary conditions and source to set up the stage to apply physical-informed-machine-learning (PIML) with neural networks to solve the boundary value problem.

### 2.1. Model formulation

We formulate the groundwater flow model in a general time-dependent setting. We consider flow of ground water in a given domain  $\mathcal{D}$  with piecewise smooth boundary  $\partial\mathcal{D}$ , in which some parts are impervious. We denote the water potential by  $h(\mathbf{x}, t)$  at location  $\mathbf{x}$  and time  $t$ . It is related to the hydrostatic pressure through

$$h(\mathbf{x}, t) = \frac{p - p_0}{\rho g} \quad (2.1)$$

where  $g$  is the gravity acceleration,  $\gamma$  is the height of the water basin measured from the bottom impervious layer,  $\rho(p)$  is the density of water, a function of pressure  $p$ ,  $p_0$  is the atmospheric pressure at the top surface of the water table and  $p(\mathbf{x}, t)$  is the hydrostatic pressure at  $\mathbf{x}$ . Since the water potential is a gauge variable, we choose the origin of the coordinate system at the lower impervious layer so that the water potential at the surface is determined by the altitude of the top surface relative to the impervious layer. We remark that the origin for  $\gamma$  is chosen as the lowest point along the bottom surface when it is not flat. The flow equation of  $h(\mathbf{x}, t)$  is given by the following continuity equation:

$$\frac{\partial h}{\partial t}$$

$$\frac{\partial S}{\partial t} = \nabla \cdot \mathbf{v} + Q, \quad (2.2)$$

where  $S$  is the storage rate,  $Q$  is the source term,  $\mathbf{v}$  is the effective velocity or the Darcy velocity. It follows from (2.1) that

$$\nabla h = \frac{1}{\rho} (\nabla p + \nabla \gamma), \quad (2.3)$$

The constitutive equation between water potential  $h$  and Darcy velocity  $\mathbf{v}$  is given by the Darcy's law (Nield and Bejan, 2006):

$$\mathbf{v} = K \cdot \nabla h, \quad (2.4)$$

where  $K$  is the mobility or conductivity coefficient tensor. We note that (2.4) can be viewed as a force balance equation, where the inverse,  $K^{-1}$ , serves as the friction coefficient. It follows from (2.2) and (2.4) that

$$\frac{\partial h}{\partial t} = \nabla \cdot (K \cdot \nabla h) + Q. \quad (2.5)$$

This is the governing equation for water potential  $h$  from which the Darcy velocity is inferred.

## 2.2. Dirichlet boundary-value problem

In a water basin  $\mathcal{D}$ , this partial differential equation is accompanied by a set of boundary conditions over domain boundary  $\Gamma = \partial\mathcal{D}$ . We consider the following 2D domain with boundary conditions given below (see Fig. 2.1),

$$\mathbf{n} \cdot K \cdot \nabla h|_{\Gamma_{b,l,r}} = 0, \quad h(x, \phi(x))|_{\Gamma_t} = g\phi(x), \quad (2.6)$$

where  $\mathbf{n}$  is the unit external normal to the boundary,  $\Gamma_{b,l,r,t}$  are the boundaries at the bottom, left, right and top side of domain  $\mathcal{D}$ , respectively, and equation  $y = \phi(x)$  defines the top boundary ( $\Gamma_t$ ). The lateral boundaries are assumed vertical line segments in domain  $\mathcal{D}$  while the top and bottom ones can be arbitrary. We name this domain the Toth basin for its origin in the Toth's seminal paper on the Toth water table. Notice that the lateral boundaries and the bottom one are assumed impervious in the Toth basin while the top one is not (Toth, 1963). When the bottom boundary is flat and top boundary inclined with a small slope, Toth calculated his well-known Toth water table solution in Toth (1970) using an approximate analytical method based on an asymptotic analysis on a rectangular domain.

Given any boundary conditions along  $\partial\mathcal{D}$ , we need to check their consistence with the governing equation in  $\mathcal{D}$  (Li et al., 2021b). We integrate Eq. (2.5) over  $\mathcal{D}$  to obtain

$$\int_{\mathcal{D}} \frac{\partial h}{\partial t} d\mathbf{x} - \int_{\mathcal{D}} \nabla \cdot (K \cdot \nabla h) d\mathbf{x} + \int_{\mathcal{D}} Q d\mathbf{x} = \int_{\Gamma} [S \frac{\partial h}{\partial t} - Q] d\mathbf{x} - \int_{\Gamma} \mathbf{n} \cdot (K \cdot \nabla h) dS = 0. \quad (2.7)$$

It imposes a consistent condition between the boundary conditions on  $h$  and the solution in the interior. If boundary conditions are given in (2.6), the consistent condition reduces to

$$\int_{\mathcal{D}} \frac{\partial h}{\partial t} d\mathbf{x} - \int_{\Gamma_t} Q d\mathbf{x} - \int_{\Gamma_t} \mathbf{n} \cdot (K \cdot \nabla h) dS = 0. \quad (2.8)$$

In steady states and without the source term, in particular, the consistent condition further reduces to

$$\int_{\Gamma_t} \mathbf{n} \cdot (K \cdot \nabla h) dS = 0. \quad (2.9)$$

$\Gamma_t$

The consistent condition is a crucial constraint for the equation to have a steady state solution. Physically, this condition indicates that the netflux across the top boundary in steady state must be zero. For the given top boundary  $\gamma = \phi(x)$ , the unit external normal  $\mathbf{n}$  couples  $\phi$  to solution  $h(\mathbf{x})$  obtained in  $\mathcal{D}$  through (2.9).

We summarize the mixed boundary value problem with the Dirichlet boundary condition on the top as follows

$$\begin{cases} \nabla \cdot K \cdot \nabla h = 0, & \mathbf{x} \in \mathcal{D}, \\ \mathbf{n} \cdot K \cdot \nabla h|_{\Gamma_{b,l,r}} = 0, & h(x, \phi(x))|_{\Gamma_t} = g\phi(x). \end{cases} \quad (2.10)$$

Assuming the boundary-value problem is well-posed,  $h$  is a solution of (2.10), and  $\hat{h}$  is another function of the same regularity as  $h$ ,  $\hat{h}$  satisfies the following estimate:

$$\|h - \hat{h}\|_{\mathcal{D}} \leq C_1 \|\nabla \cdot K \cdot \nabla h\|_{\mathcal{D}} + C_2 \|\mathbf{n} \cdot K \cdot \nabla h\|_{\Gamma_{b,l,r}} + C_3 \|h - g\phi\|_{\Gamma_t}, \quad (2.11)$$

$$\|h - \hat{h}\|_{\mathcal{D}} \leq C_1 \|\nabla \cdot K \cdot \nabla h\|_{\mathcal{D}} + C_2 \|\mathbf{n} \cdot K \cdot \nabla h\|_{\Gamma_{b,l,r}} + C_3 \|h - g\phi\|_{\Gamma_t}. \quad (2.12)$$

$$Loss = C_1 \|\nabla \cdot K \cdot \nabla h\|_{\mathcal{D}} + C_2 \|\mathbf{n} \cdot K \cdot \nabla h\|_{\Gamma_{b,l,r}} + C_3 \|h - g\phi\|_{\Gamma_t}. \quad (2.13)$$

In this case, finding the solution of (2.10) is turned into a minimization problem of the residues in (2.13). This is the foundation of PIML formulation (Karniadakis et al., 2021). The crucially important part in this formulation is the choice of the norms in the loss function so that it is consistent with the well-posedness proof of the initial-boundary value problem (Raissi et al., 2019). In practice, we augment the loss function defined in (2.13) by a penalization of the consistent condition given in (2.9) as follows

$$Loss = C_1 \|\nabla \cdot K \cdot \nabla h\|_{\mathcal{D}} + C_2 \|\mathbf{n} \cdot K \cdot \nabla h\|_{\Gamma_{b,l,r}} + C_3 \|h - g\phi\|_{\Gamma_t} + L_1 \left( \int_{\Gamma_t} \mathbf{n} \cdot (K \cdot \nabla h) dS \right)^2, \quad (2.14)$$

where  $L_1 > 0$  is a model parameter set by the user. In this paper, we set  $L_1 = 1$ .

## 2.3. Robin boundary-value problem

A more physical boundary condition in steady states of the groundwater flow equation at the top boundary perhaps should be

$$\mathbf{n} \cdot K \cdot \nabla h = -\gamma(h - g\phi(x)), \quad (2.15)$$

where  $\gamma$  is the rate parameter. It indicates that the flux through the top boundary is proportional to the difference of the water potential and the saturated steady state water potential. If  $\gamma = 0$ , (2.15) reduces to the impervious Neumann boundary condition; whereas it reduces to the Dirichlet one if  $\gamma \rightarrow \infty$ .

If we assume that the Robin boundary-value problem is well-posed,  $h$  is a solution, and  $\hat{h}$  a function in  $H^1(\mathcal{D})$ , it follows from the wellposedness that

$$\|h - \hat{h}\| \leq C_1 \|\nabla \cdot K \cdot \nabla h\|_{\mathcal{D}} + C_2 \|\mathbf{n} \cdot K \cdot \nabla h\|_{\Gamma_{b,l,r}} + C_3 \|\mathbf{n} \cdot K \cdot \nabla h + \gamma(h - g\phi(x))\|_{\Gamma_t}, \quad (2.16)$$

where  $C_i, i=1,2,3$  are positive constants. The loss function can then be devised as follows

$$Loss = C_1 \|\nabla \cdot K \cdot \nabla h\|_{\mathcal{D}} + C_2 \|\mathbf{n} \cdot K \cdot \nabla h\|_{\Gamma_t - \Gamma_b} + C_3 \|\mathbf{n} \cdot K \cdot \nabla h + \gamma(h - g\phi(x))\|_{\Gamma_t}. \quad (2.17)$$

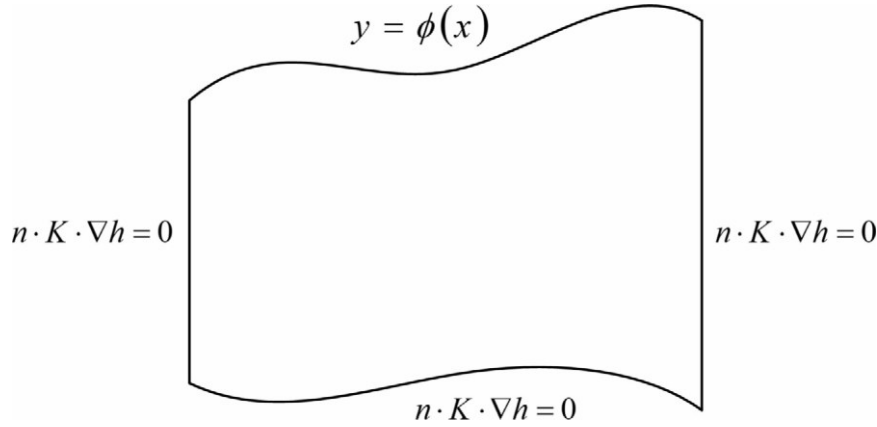
This loss function penalizes all the residues in the equation and the boundary conditions.

The steady state governing equation without the source together with boundary condition (2.15) yields the following consistency condition:

$$\int_{\Gamma_t} \gamma(h - \phi) ds = 0. \quad (2.18)$$

(2.9) and (2.18) are two constraints for the solution to satisfy the Dirichlet and the Robin boundary condition, respectively, which must be ensured in any solution solvers. In practice, the loss function used in machine-learning in this study is given by

$$Loss = C_1 \|\nabla \cdot K \cdot \nabla h\|_{\mathcal{D}} + C_2 \|\mathbf{n} \cdot K \cdot \nabla h\|_{\Gamma_t - \Gamma_b} + C_3 \|\mathbf{n} \cdot K \cdot \nabla h + \gamma(h - g\phi(x))\|_{\Gamma_t} + \int_{\Gamma_t} \gamma(h - \phi) ds)^2. \quad (2.19)$$



**Fig. 2.1.** Toth basin  $\mathcal{D}$  and the prescribed boundary conditions over  $\partial\mathcal{D}$  embedded in a rectangular domain  $[0,1] \times [0,1]$ . The top and bottom boundaries are given by  $y = \phi(x)$  and  $y = \phi_2(x)$ , respectively.

## 2.4. Nondimensionalization

In order to solve the equations together with the boundary conditions numerically, we need to nondimensionalize them. We introduce length scale in  $x$ :  $L_x$ , in  $y$ :  $L_y$ , and time scale:  $T$ , respectively. The dimensionless variables are defined as follows

$$\tilde{x} = x/L_x, \tilde{y} = y/L_y, \tilde{t} = t/T, \tilde{h} = h/h_0, \tilde{\phi} = \phi/h_0, \tilde{Q} = Q/(S_0 h_0) \quad (2.20)$$

where  $h_0$  is a characteristic water potential. The top Dirichlet boundary condition is given by

$$\tilde{h} = \tilde{\phi}. \quad (2.21)$$

We denote the characteristic storage rate by  $S_0$ . The dimensionless model parameters are given by

$$\tilde{S} = \frac{S}{S_0}, \tilde{K} = \frac{K}{L_{2y} S_0}, \tilde{Q} = \frac{Q}{S_0 h_0} \quad (2.22)$$

where

$$\mathbf{A} = \begin{pmatrix} \epsilon & 0 \\ 0 & 1 \end{pmatrix}, \quad \epsilon = \frac{L_x}{L_y} \quad (2.23)$$

is the aspect ratio of the basin. The flow equation in the dimensionless form is given by

$$\tilde{S} \frac{\partial \tilde{h}}{\partial \tilde{t}} = \nabla \cdot \tilde{K} \cdot \nabla \tilde{h} + \tilde{Q}. \quad (2.24)$$

$$\text{We choose } \tilde{h}_0 = g L_y, \quad \tilde{\partial \tilde{h}} \sim \sim \sim \quad (2.25)$$

and drop the  $\tilde{\cdot}$  from the dimensionless equations to obtain the dimensionless equation and boundary conditions as follows:

$$\begin{cases} S \frac{\partial h}{\partial t} = \nabla \cdot K \cdot \nabla h + Q, \\ h = \phi(x), \mathbf{x} \in \Gamma_t, \mathbf{n} \cdot K \cdot \nabla h = 0, \mathbf{x} \in \partial\mathcal{D} - \Gamma_t, \end{cases} \quad (2.26)$$

The consistent condition (2.9) retains.

Analogously, we obtain the dimensionless Robin boundary condition at the top boundary as follows

$$\mathbf{n} \cdot K \cdot \nabla h = -\tilde{\gamma}(h - \phi), \quad (2.27)$$

where  $\tilde{\gamma} = \gamma L_x$ . We drop the tilde over  $\gamma$  for brevity in the following. In this paper, we consider  $K = \text{Diag}(K_{11}, K_{22})$  as a diagonal mobility matrix in the dimensionless equation,  $x \in [0,1]$ , and  $0 < \phi_2(x) < \phi(x)$

1 as bottom and top boundaries, where  $y = \phi_2$  represents the bottom boundary.

Next, we present three implementations of the DeepONet for the mapping from specified boundaries and the boundary conditions (2.6) to the steady state solution of the groundwater flow equation in  $\mathcal{D}$  (Lu et al., 2021), from which the Darcy's velocity can be recovered.

## 3. DeepONet for the boundary-to-solution mapping

For the boundary value problem in the Toth basin, we would like to establish a mapping from the boundary and the associated boundary condition to the solution of the steady state groundwater flow equation in the domain. We adopt the physics-informed machine learning approach and use the

DeepONet as the neural network to represent the mapping (Lu et al., 2021). For the Toth basin, we consider a domain with flat lateral, arbitrary top and bottom boundaries. Firstly, we consider a domain of aspect ratio  $\epsilon$  with flat and impervious bottom and lateral boundaries and prescribed water potential at the top boundary,  $y = \phi(x)$ . We construct the top surface to the solution mapping in the Toth basin with  $(\phi, \epsilon)$  as the input. Owing to the fact that the dimensionless boundary condition coincides with the boundary representation, we only need to learn a mapping from top boundary  $y = \phi(x)$  with aspect ratio  $\epsilon$  to the solution in  $\mathcal{D}$ . We present three different approaches to accomplishing this goal using two distinct representations of top boundary  $y = \phi(x)$ , respectively. Secondly, we discuss an extension of the approach to the domain where the bottom boundary and the top boundary are both arbitrary.

### 3.1. Piecewise polynomial interpolation of the top boundary

We represent top boundary  $y = \phi(x)$  using  $n$  discrete points  $\mathbf{x}_i = (x_i, \phi(x_i))$ ,  $i = 1, \dots, n$  uniformly distributed in the  $x$ -coordinate, where  $x_i = (i-1)\Delta x$ ,  $\Delta x = \frac{1}{n-1}$ . We acknowledge a new development in treating the boundary condition in a weak formulation of partial differential equations by introducing new variable to satisfy the homogeneous boundary conditions in PIML (Sukumar and Srivastava, 2022). However, our proposed approach suffices for the current problem. We denote the approximate solution of this mixed boundary value problem in the interior of  $\Omega$  by a DeepONet  $\mathcal{Q}(\mathbf{h}, \mathbf{c}; \mathbf{x})$  as follows

$$\mathcal{Q}(\mathbf{h}_t, \epsilon, \mathbf{x}) = \sum_{k=1}^p \sum_{i=1}^q c_i^k \sigma \left( \sum_{j=1}^n \xi_{ijk} \phi(\mathbf{x}_j) + \xi_{i0}^k \epsilon + \theta_i^k \right) \sigma(W_k \cdot \mathbf{x} + \zeta_k) + b_0, \quad (3.1)$$

where  $c_k^i, \xi_{ij}^k, W_k$  are weights and  $\theta_k^i, \zeta_{ij}^k, b_0$  are biases of the neural network,  $\mathbf{h}_t = (\phi(x_1), \dots, \phi(x_n)) \in \mathbb{R}^n$  denotes the uniformly distributed,  $y$ -coordinates of the interpolating points at the top boundary, and  $n, p, q$  are positive integers. The DeepONet represents the mapping from  $\mathbf{h}_t \in \mathbb{R}^n$  to the solution.

To apply the PIML method to learn the neural network, we choose  $n_b, n_r, n_l, n_b, n_r, n_t$  points at the left, right, bottom, and top boundary, and the interior randomly. For convenience, we use odd number for  $n_i = n$  at the top boundary. The loss function of the machine learning model is given by (2.14) with the  $L_2$  norms in the interior and on the boundary. We evaluate the integral norms using the Monte Carlo sampling. For the randomly chosen points along the boundary and in the interior,  $\{\mathbf{x}_i, i = l, r, b, t\}$ , and a well-defined uniform division of  $[0, 1]$ ,  $\{x_j, j = 1, \dots, n_i\}$ ,  $\mathbf{x}_j = (x_j, \phi(x_j))$ , the specific expression of the loss function is given by

$$\begin{aligned} L(\mathcal{Q}_{\mathbf{h}_\ell}^i, \epsilon) &= \sum_{i=1}^n (\phi(\mathbf{x}_i) - G(\mathbf{h}_\ell, \epsilon; \mathbf{x}_i))^2 + \frac{1}{n} \sum_{j=1}^{n_i} (\nabla \cdot \mathbf{K} \cdot \nabla G \\ &\quad \left( \frac{\partial}{\partial \mathbf{x}} \left( \mathcal{G}(\mathbf{h}_\ell, \epsilon; \mathbf{x}_i) \right)^2 + \frac{1}{n} \sum_{i=1}^{n_i} \left( \frac{\partial}{\partial \mathbf{x}} \mathcal{G}(\mathbf{h}_\ell, \epsilon; \mathbf{x}_i) \right)^2 + \frac{1}{n} \sum_{i=1}^{n_i} \left( \frac{\partial}{\partial \mathbf{y}} \mathcal{G}(\mathbf{h}_\ell, \epsilon; \mathbf{x}_i) \right)^2 \right) \\ &\quad + \sum_{i \in \text{even} \in \{2, \dots, n-1\}} F(\mathbf{h}_\ell, \epsilon; \mathbf{x}_i) + \sum_{i \in \text{odd} \in \{3, \dots, n-2\}} F(\mathbf{h}_\ell, \epsilon; \mathbf{x}_i) + F(\mathbf{h}_\ell, \epsilon; \mathbf{x}_{nt}) \end{aligned} \quad (3.2)$$

where the boundary mass flux is given by

$$A(\mathbf{h}_\phi, \epsilon, \mathbf{x}) = \mathbf{n}^{(\mathcal{A})} \cdot K \cdot \left( \frac{\partial}{\partial x_1} G(\mathbf{h}_\phi, \epsilon, \mathbf{x}), \frac{\partial}{\partial x_2} G(\mathbf{h}_\phi, \epsilon, \mathbf{x}) \right)^T = -K_{11} \phi_i(\mathbf{x}) G_i(\mathbf{h}_\phi, \epsilon, \mathbf{x}) + K_{22} G_j(\mathbf{h}_\phi, \epsilon, \mathbf{x}). \quad (3.3)$$

The Simpson’s quadrature formula is employed (Butcher, 2016) to ensure the integral is accurate up to the fourth order in  $\Delta x$ . This is the PIML formulation of the problem where the residues in the equation and boundary conditions are penalized in the loss function in the  $L_2$  norm. This loss function is defined for each given top boundary parameterized by  $\mathbf{h}_x$  and a set of randomly selected points from other parts of the domain. We note that this loss also includes a penalization term for constraint (2.9) to enforce consistency.

In the practical implementation, we modify the loss function by rebalancing the weights. The loss function used in machine-learning is then modified into

$$\begin{aligned}
\mathcal{L}(\theta, \mathbf{h}_t, \epsilon) &= \frac{\lambda_1}{n} \sum_{i=1}^n (\phi(\mathbf{x}_i) - G(\mathbf{h}_t, \epsilon, \mathbf{x}_i))^2 \\
&+ \frac{\lambda_2}{n} \sum_i (\nabla \cdot K \cdot \nabla G_{n \neq 1}(\mathbf{h}_t, \epsilon, \mathbf{x}_i))^2 + \\
&\quad \frac{\lambda_3}{n} \sum_{i=1}^{n_l} \left( \frac{\partial}{\partial x} \right)^2 + \frac{\lambda_4}{n} \sum_{i=1}^{n_r} \left( \frac{\partial}{\partial y} \right)^2 \mathcal{G}(\mathbf{h}_t, \epsilon, \mathbf{x}_i) \quad \left( \frac{\partial}{\partial x} \mathcal{G}(\mathbf{h}_t, \epsilon, \mathbf{x}_i) \right)^2 \\
&+ \frac{\lambda_5}{n_b} \sum_{i=1}^{n_b} \left( 10 \times \frac{\partial}{\partial y} \right)^2 \mathcal{G}(\mathbf{h}_t, \epsilon, \mathbf{x}_i) \quad \left( \frac{\partial}{\partial x} \mathcal{G}(\mathbf{h}_t, \epsilon, \mathbf{x}_i) \right)^2 \quad (3.4)
\end{aligned}$$

where the weights are re-balanced as follows in each iteration

$$\begin{aligned}
loss_{stop} &= \frac{1}{n} \sum_{i=1}^n (\phi(\mathbf{x}_i) - G(\mathbf{h}_t, \epsilon, \mathbf{x}_i))^2, \\
loss_{seq} &= \frac{1}{n} \sum_{i=1}^{n_i} (\nabla \cdot K \cdot \nabla G(\mathbf{h}_t, \epsilon, \mathbf{x}_i))^2, \\
loss_{left} &= \frac{1}{n} \sum_{i=1}^{n-1} (G(\mathbf{h}_t, \epsilon, \mathbf{x}_i))^2, \quad loss_{right} = \frac{1}{n} \sum_{i=1}^{n-1} (G(\mathbf{h}_t, \epsilon, \mathbf{x}_{i+1}))^2, \\
loss_{bottom} &= \frac{1}{n_b} \sum_{b=1}^{n_b} (G(\mathbf{h}_t, \epsilon, \mathbf{x}_b))^2, \\
loss_{con} &= \frac{1}{3} \sum_{i \in \{2, \dots, n-1\}} F(\mathbf{h}_t, \epsilon, \mathbf{x}_i) + \frac{1}{4} \sum_{i \in \{2, \dots, n-1\}} F(\mathbf{h}_t, \epsilon, \mathbf{x}_{i+1}) \\
&\quad + \frac{1}{2} \sum_{i \in \{3, \dots, n-2\}} F(\mathbf{h}_t, \epsilon, \mathbf{x}_i) \\
&\quad + \frac{1}{2} \sum_{i \in \{3, \dots, n-2\}} F(\mathbf{h}_t, \epsilon, \mathbf{x}_{i+1}) \\
loss_{i-1} &= (loss_{stop}^{i-1} + loss_{seq}^{i-1} + loss_{right}^{i-1} + loss_{left}^{i-1} + \\
loss_{bottom}^{i-1} &\quad + loss_{con}^{i-1}) / 6, \\
\lambda_1 &= loss_{stop}^{i-1} / loss_{seq}^{i-1}, \quad \lambda_2 = loss_{seq}^{i-1} / loss_{right}^{i-1}, \quad \lambda_3 = loss_{right}^{i-1} / loss_{left}^{i-1}, \\
\lambda_4 &= loss_{left}^{i-1} / loss_{bottom}^{i-1}, \quad \lambda_5 = loss_{bottom}^{i-1} / loss_{con}^{i-1}, \quad \lambda_6 = loss_{con}^{i-1} / loss_{seq}^{i-1}
\end{aligned}
\tag{3.5}$$

For a given set of randomly chosen top boundary dataset in  $\mathbf{h}_\ell: \mathbf{h}_\ell^{(1)}, \dots, \mathbf{h}_\ell^{(m)}$ , and the aspect ratio  $\epsilon_\ell^{(l)}, l = 1, \dots, L$ , we define the total loss function as follows

$$\mathcal{L}(\theta) = \frac{1}{m} \sum_{\ell=1}^m \sum_{j=1}^m \mathbb{E}_{\mathbf{h}_{\ell,j} \in \mathcal{H}} [\ell(\theta; \mathbf{h}_{\ell,j})]. \quad (3.6)$$

We remark that the numbers of randomly selected interior and boundary points at each given  $\mathbf{h}_i$  and  $\mathcal{A}^d$  are not the same so that  $\mathcal{L}(\theta, \mathbf{h}^d, \mathcal{A}^d)$  can have different number of terms in the sums. We point it out that choices of activation functions are important to the performance of machine learning

model. In this model, we use tanh as the activation function. If one uses DeepONet to solve non-linear equations, it is better off to use smooth activation functions. Our experience with ReLU for this problem is not as good as the one using the tanh function as the activation function.

### 3.2. Spectral representation of the boundary

Alternatively, we represent continuous top boundary  $\phi(x)$  using a truncated Sine Fourier series together with a linear interpolation function as follows (Haberman, 2013):

$$\phi(x) = \phi(0) + \frac{\phi(L) - \phi(0)}{L}x + \sum_{j=1}^m \frac{j\pi}{L} b_j \sin \frac{j\pi}{L}x, \quad (3.7)$$

where  $m$  is the number of modes in the spectral expansion and  $b_j$  is the  $j$ th Sine Fourier coefficient given by

$$b_j = -\int_0^L [\phi(x) - \phi(0) - \frac{\phi(L) - \phi(0)}{L}x] \sin(\frac{j\pi}{L}x) dx. \quad (3.8)$$

We represent the top boundary using  $m+2$  discrete values  $\mathbf{h}_t = (\phi(0), \phi(L), b_1, \dots, b_m) \in \mathbb{R}^{m+2}$ , consisting of the Sine Fourier coefficients and the two end point values. Given the boundary condition at top boundary  $\phi(\mathbf{x})$ , we want to learn a mapping from  $(\mathbf{h}_t, \epsilon)$  to the solution of the steady state governing equation in  $\Omega$ .

We denote the solution of the boundary value problem in the interior of  $\Omega$  by DeepONet  $G(\mathbf{h}_t, \mathbf{x})$  as follows:

$$G(\mathbf{h}_t, \epsilon, \mathbf{x}) = \sum_{k=1}^p \sum_{i=1}^q c_i^k \sigma \left( \sum_{j=1}^{m+2} \xi_{ij}^k \mathbf{h}_{t,j} + \xi_{i0}^k \epsilon + \theta_i^k \right) \sigma(W_k \cdot \mathbf{x} + \zeta_k) + b_0, \quad (3.9)$$

where  $c_i^k, \xi_{ij}^k, W_k$  are weights and  $\theta_i^k, \zeta_k, b_0$  biases. We randomly sample  $n_l, n_r, n_b$  points from the left, right and bottom boundary respectively,  $n_i$  points in the interior of  $\Omega$ . We divide  $[0, 1]$  uniformly into  $n-1$  intervals, separated by  $x_i = (i-1)\Delta x, \Delta x = \frac{1}{n-1}, i = 1, \dots, n$ . We use the DeepONet to learn the mapping from  $(\mathbf{h}_t, \epsilon)$  to solution  $h(\mathbf{x}, t)$  in  $\Omega$ .

The cost function in the model for each given top boundary, the randomly chosen points along the boundary and in the interior,  $\{\mathbf{x}^i, i = l, r, b, \mathbf{x}\}$ , a well-defined uniformed division of  $[0, 1]$ ,  $\{x_j, j = 1, \dots, n\}$ , that defines the top boundary points  $\mathbf{x}^f = (x_j, \phi(x_j))$ , and aspect ratio  $\epsilon$  is then defined by

$$\begin{aligned} \mathcal{L}(\theta, \mathbf{h}_t, \epsilon) = & \frac{1}{n} \sum_i (\nabla \cdot \mathbf{K} \cdot \nabla G_{j=1}(\mathbf{h}_t, \epsilon, \mathbf{x}_i))^2 + \frac{1}{n} \sum_{i=1}^{n_l} (G_x(\mathbf{h}_t, \epsilon, \mathbf{x}_i))^2 + \frac{1}{n} \sum_{i=1}^{n_r} (G_y(\mathbf{h}_t, \epsilon, \mathbf{x}_i))^2 \\ & + \frac{1}{n} \sum_{i=1}^{n_b} (G_z(\mathbf{h}_t, \epsilon, \mathbf{x}_i))^2 + \frac{1}{n} \sum_{i=1}^{n_i} (G(\mathbf{h}_t, \epsilon, \mathbf{x}_i) - \phi(\mathbf{x}_i))^2 \\ & + \frac{1}{3} \sum_{i \in \{1, \dots, n-1\}} F(\mathbf{x}_i, \epsilon) + \frac{1}{4} \sum_{i \in \{2, \dots, n-1\}} F(\mathbf{x}_i, \epsilon) \\ & + \frac{1}{2} \sum_{i \in \{3, \dots, n-2\}} F(\mathbf{x}_i, \epsilon) + F(\mathbf{x}_n, \epsilon) \end{aligned} \quad (3.10)$$

In the practical implementation, we once again re-balance the “local loss” as alluded to earlier.

For the bounded Toth basin with two vertical, lateral boundaries, we can rescale or transform the bounded, arbitrary physical domain into a rectangular domain and then solve the equation in the rectangular domain. We call this the domain mapping approach.

### 3.3. Domain mapping

We present yet another alternative approach to establish the mapping from the top boundary to the solution in a Toth basin using a nonlinear domain mapping. We assume the top boundary is given by  $y = \phi(x) > 0$  and the bottom one by  $y = 0$  for  $x \in [0, L]$ . We introduce a change of variable from  $(x, y)$  to  $(x, z)$  as follows

$$x = x, \quad z = \frac{y}{\phi(x)}, \quad y \in [0, \phi(x)]. \quad (3.11)$$

The 2D gradient operator in the new coordinate is given by

$$\nabla^* = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial z} \right) = \left( \frac{\partial}{\partial x} + \frac{y}{\phi(x)} \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right). \quad (3.12)$$

Or equivalently,

$$\nabla = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right) = \left( \frac{\partial}{\partial x} + \frac{y}{\phi(x)} \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right) \begin{pmatrix} \frac{\partial x}{\partial x} & \frac{\partial y}{\partial x} \\ \frac{\partial x}{\partial y} & \frac{\partial y}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & -\frac{y}{\phi^2(x)} \\ 0 & \frac{1}{\phi(x)} \end{pmatrix} \nabla^*. \quad (3.13)$$

We denote

$$D = \begin{pmatrix} 1 & -\frac{y}{\phi^2(x)} \\ 0 & \frac{1}{\phi(x)} \end{pmatrix} = \begin{pmatrix} 1 & -\frac{y}{\phi^2(x)} \\ 0 & \frac{1}{\phi(x)} \end{pmatrix}. \quad (3.14)$$

The Laplace equation is rewritten into

$$(D \cdot \nabla^*) \cdot K \cdot (D \cdot \nabla^*) h = 0. \quad (3.15)$$

The boundary conditions of  $h$  is given by

$$\begin{cases} \mathbf{n} \cdot K \cdot \nabla h|_{\Gamma_{L,R}} = \mathbf{n} \cdot K \cdot (D \cdot \nabla^*) h|_{\Gamma_{L,R}} = 0, \\ h(x, 1)|_{\Gamma_{top}} = \phi(x) \quad (\text{Dirichlet}), \\ \mathbf{n} \cdot K \cdot (D \cdot \nabla^*) h|_{\Gamma_{top}} = -\gamma(h - \phi) \quad (\text{Robin}), \quad x \in [0, 1]. \end{cases} \quad (3.16)$$

In this study, we limit ourselves to

$$K = \text{diag}(K_{11}, K_{22}). \quad (3.17)$$

Then,

$$\mathbf{n} \cdot K \cdot (D \cdot \nabla^*) h|_{\Gamma_{L,R}} = K_{22} \frac{\partial h}{\partial z} \bigg|_{\Gamma_{L,R}} = 0, \quad \mathbf{n} \cdot K \cdot (D \cdot \nabla^*) h|_{\Gamma_{top}} = -\gamma(h - \phi). \quad (3.18)$$

These imply

$$\frac{\partial h}{\partial x} \bigg|_{\Gamma_{L,R}} = 0, \quad \frac{\partial h}{\partial z} \bigg|_{\Gamma_{L,R}} = 0, \quad \frac{\partial h}{\partial x} \bigg|_{\Gamma_{top}} = 0, \quad \frac{\partial h}{\partial z} \bigg|_{\Gamma_{top}} = 0, \quad (3.19)$$

where the bottom boundary is assumed flat.

The steady state governing equation without a source is given by

$$\begin{aligned} (D \cdot \nabla^*) \cdot K \cdot (D \cdot \nabla^*) h \\ = K_{11} \left[ \frac{\partial^2 h}{\partial x^2} - \frac{\partial}{\partial x} \left( -\frac{y}{\phi^2(x)} \frac{\partial h}{\partial x} \right) - \frac{y}{\phi^2(x)} \frac{\partial^2 h}{\partial x^2} \right] + K_{22} \frac{\partial^2 h}{\partial z^2} = 0. \end{aligned} \quad (3.20)$$

The consistency condition becomes

$$\int_{z=1}^{\infty} [\mathbf{n}^{(d)} \cdot \mathbf{K} \cdot (\mathbf{D} \cdot \nabla^*) / h] dx = \int_{z=1}^{\infty} [K_{22} - K_{11} \phi_x (h_x - \phi_{xx} h_x)] dx = 0, \quad (3.21)$$

where  $\mathbf{n}^{(d)} = (-\phi_x, 1)$ . We represent the top boundary using  $m+2$  discrete values  $\mathbf{h}_x = (\phi(a), \phi(b), b_x, \dots, b_m)$  from the truncated Sine Fourier series approximation. Given the boundary condition at the top boundary  $z = 1$ , we want to learn a mapping from  $\mathbf{h}_x$  to the solution of the steady state governing equation in  $\mathcal{Q}$ .

We denote the solution in the interior of  $\mathcal{Q}$  by DeepONet  $\mathcal{G}(\mathbf{h}_x, \mathbf{x})$  defined in (3.9). The loss function for each given top boundary  $\mathbf{h}_x$ , aspect ratio  $\epsilon$ , the randomly chosen points along the boundary and in the interior,  $\{\mathbf{x}_i, i = 1, \dots, n_b\}$ , and a well-defined uniform division of  $[0, 1]$ ,  $\{x_j\}$ , that defines the top boundary points  $\mathbf{x}_f = (x_j, 1)$ , is then defined by

$$\begin{aligned} \mathcal{L}(\theta, \mathbf{h}_x, \epsilon) = & n^{-1} \sum_{i=1}^n [(D \cdot \nabla^*) \cdot \mathbf{K} \cdot (D \cdot \nabla^*) \mathcal{G}(\mathbf{h}_x, \mathbf{x}_i)]^2 \\ & + \frac{1}{n} \sum_{j=1}^{n_t} [\phi(\mathbf{x}_j^t) - \mathcal{G}(\mathbf{h}_x, \mathbf{x}_j^t)]^2 \\ & + \frac{1}{n} \sum_{i=1}^n [\mathbf{n} \cdot \mathbf{K} \cdot (D \cdot \nabla^*) \mathcal{G}(\mathbf{h}_x, \mathbf{x}_i)]^2 + \sum_{i=1}^n [K_{11}(\mathbf{h}_x, \epsilon, \mathbf{x}_i)]^2 \\ & + \frac{1}{n} \sum_{i=1}^n [\mathbf{n} \cdot \mathbf{K} \cdot (D \cdot \nabla^*) \mathcal{G}(\mathbf{h}_x, \mathbf{x}_i)]^2 \end{aligned} \quad (3.22)$$

$$\begin{aligned} & + \frac{1}{n} \sum_{i=1}^n [\mathbf{n} \cdot \mathbf{K} \cdot (D \cdot \nabla^*) \mathcal{G}(\mathbf{h}_x, \mathbf{x}_i)]^2 + \frac{1}{n} \sum_{i=1}^n [F(\mathbf{h}_x, \epsilon, \mathbf{x}_i)]^2 \\ & + \frac{1}{n} \sum_{i=1}^n [F(\mathbf{h}_x, \epsilon, \mathbf{x}_i)]^2 + \frac{1}{n} \sum_{i=1}^n [F(\mathbf{h}_x, \epsilon, \mathbf{x}_i)]^2 \end{aligned}$$

where

$$F(\mathbf{x}, \epsilon) = K_{22} - K_{11} \phi_x (h_x - \phi_{xx} h_x) - G_x. \quad (3.23)$$

In the practical implementation, we adopt a re-balanced or modified loss function, in which we add a weight to each term in the loss. The modified loss function is given by

$$\begin{aligned} \mathcal{L}(\theta, \mathbf{h}_x, \epsilon) = & \lambda n^{-1} \sum_{i=1}^n [\phi(\mathbf{x}_i^t) - \mathcal{G}(\mathbf{h}_x, \mathbf{x}_i^t)]^2 \\ & + \lambda n^{-2} \sum_{i=1}^n [(D \cdot \nabla^*) \cdot \mathbf{K} \cdot (D \cdot \nabla^*) \mathcal{G}(\mathbf{h}_x, \mathbf{x}_i)]^2 \\ & + \frac{\lambda_3}{n} \sum_{i=1}^{n_t} [G_x(\mathbf{h}_x, \epsilon, \mathbf{x}_i)]^2 + \frac{\lambda_4}{n} \sum_{i=1}^{n_r} [G_x(\mathbf{h}_x, \epsilon, \mathbf{x}_i)]^2 \end{aligned} \quad (3.24)$$

$$+ \lambda n^{-2} \sum_{i=1}^n [10 \times \mathbf{n} \cdot \mathbf{K} \cdot (D \cdot \nabla^*) \mathcal{G}(\mathbf{h}_x, \epsilon, \mathbf{x}_i)]^2$$

$$+ \lambda_6 \left[ \frac{\Delta x}{3} (F(\mathbf{x}_1^t, \epsilon) + 4 \sum_{i \text{ even} \in \{2, \dots, n_t-1\}} F(\mathbf{x}_i^t, \epsilon) + \sum_{i \text{ odd} \in \{3, \dots, n_t-2\}} F(\mathbf{x}_i^t, \epsilon) + F(\mathbf{x}_{n_t}^t, \epsilon)) \right]^2,$$

where the weights are re-balanced as follows in each iteration

$$\begin{aligned} \text{loss}_{top} &= \frac{1}{n_t} \sum_{i=1}^{n_t} (\phi(\mathbf{x}_i^t) - \mathcal{G}(\mathbf{h}_x, \mathbf{x}_i^t))^2, \\ \text{loss}_{eq} &= \frac{1}{n_r} \sum_{i=1}^{n_r} ((D \cdot \nabla^*) \cdot \mathbf{K} \cdot (D \cdot \nabla^*) \mathcal{G}(\mathbf{h}_x, \mathbf{x}_i))^2, \\ \text{loss}_{left} &= \frac{1}{n_b} \sum_{i=1}^{n_b} (G_x(\mathbf{h}_x, \mathbf{x}_i))^2, \quad \text{loss}_{right} = \frac{1}{n_r} \sum_{i=1}^{n_r} (G_x(\mathbf{h}_x, \mathbf{x}_i))^2, \\ \text{loss}_{bottom} &= \frac{1}{n_b} \sum_{i=1}^{n_b} (G_x(\mathbf{h}_x, \mathbf{x}_i))^2, \\ \text{loss}_{con} &= \left[ \frac{\Delta x}{3} (F(\mathbf{x}_1^t, \epsilon) + 4 \sum_{i \text{ even} \in \{2, \dots, n_t-1\}} F(\mathbf{x}_i^t, \epsilon) + \sum_{i \text{ odd} \in \{3, \dots, n_t-2\}} F(\mathbf{x}_i^t, \epsilon) + F(\mathbf{x}_{n_t}^t, \epsilon)) \right]^2 \end{aligned}$$

$$\begin{aligned} & + 2 \sum_{i \text{ odd} \in \{3, \dots, n_t-2\}} F(\mathbf{x}_i^t, \epsilon) + F(\mathbf{x}_{n_t}^t, \epsilon)]^2 \\ & \text{loss}_{top} = (\text{loss}_{top}^{i-1} + \text{loss}_{eq}^{i-1} + \text{loss}_{right}^{i-1} + \text{loss}_{left}^{i-1} + \text{loss}_{bottom}^{i-1} + \text{loss}_{con}^{i-1}) / 6, \\ & \lambda^i = \frac{i-1}{100} \text{loss}_{top}^{i-1}, \quad \lambda^i = \frac{i-1}{100} \text{loss}_{eq}^{i-1}, \quad \text{loss}_{right}^i = \frac{i-1}{100} \text{loss}_{right}^{i-1}, \\ & \text{loss}_{left}^i = \frac{i-1}{100} \text{loss}_{left}^{i-1}, \quad \text{loss}_{bottom}^i = \frac{i-1}{100} \text{loss}_{bottom}^{i-1}, \quad \text{loss}_{con}^i = \frac{i-1}{100} \text{loss}_{con}^{i-1}. \end{aligned} \quad (3.25)$$

The total loss is defined in (3.6) for a given set of top boundaries. In the rescaled domain, the variable coefficient Poisson equation is solved in a rectangular domain using PIML.

### 3.4. Arbitrary bottom boundary

When the bottom boundary is varying in space as well, we denote it as  $\phi_2(x)$ . We rescale the physical domain in the  $y$  direction as follows

$$y = \frac{\phi_2(x) - \phi_1(x)}{\phi_2(x) - \phi_1(x)}, \quad y \in [\phi_2(x), \phi_1(x)]. \quad (3.26)$$

This mapping transforms the Toth basin into a rectangular domain in a new coordinate. The gradient operator is transformed as follows

$$\nabla = \left( \frac{\partial}{\partial x} - \frac{z}{\phi - \phi_2} (\phi - \phi_2)_x + \frac{\phi_{2,x}}{\phi - \phi_2} \frac{\partial}{\partial z}, \frac{1}{\phi - \phi_2} \frac{\partial}{\partial z} \right) = D \cdot \nabla^*, \quad (3.27)$$

where

$$D = \begin{pmatrix} 1 & -\frac{z}{\phi - \phi_2} (\phi - \phi_2)_x + \frac{\phi_{2,x}}{\phi - \phi_2} \\ 0 & \frac{1}{\phi - \phi_2} \end{pmatrix}. \quad (3.28)$$

The Laplace equation is rewritten into a variable coefficient one as follows

$$\begin{aligned} (D \cdot \nabla^*) \cdot \mathbf{K} \cdot (D \cdot \nabla^*) h &= 0, \quad (3.29) \text{ where} \\ (D \cdot \nabla^*) \cdot \mathbf{K} \cdot (D \cdot \nabla^*) h &= K_{11} \left[ \frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial x} \left( \frac{z}{\phi - \phi_2} (\phi - \phi_2)_x + \frac{\phi_{2,x}}{\phi - \phi_2} \right) \frac{\partial}{\partial z} \right. \\ & \left. - \left( \frac{z}{\phi - \phi_2} (\phi - \phi_2)_x + \frac{\phi_{2,x}}{\phi - \phi_2} \right) \frac{\partial}{\partial z} \right] \\ & + \frac{\phi_{2,x}}{\phi - \phi_2} \frac{\partial^2}{\partial z^2} + \left( \frac{z}{\phi - \phi_2} (\phi - \phi_2)_x + \frac{\phi_{2,x}}{\phi - \phi_2} \right) \frac{\partial}{\partial z} \left( \frac{z}{\phi - \phi_2} (\phi - \phi_2)_x + \frac{\phi_{2,x}}{\phi - \phi_2} \right) \frac{\partial}{\partial z} \\ & + K_{22} \frac{1}{(\phi - \phi_2)^2} \frac{\partial^2}{\partial z^2} = 0. \end{aligned} \quad (3.30)$$

The boundary conditions of  $h$  are given by

$$\begin{aligned} \mathbf{n} \cdot \mathbf{K} \cdot \nabla h|_{\Gamma_{L,R}} &= \mathbf{n} \cdot \mathbf{K} \cdot (D \cdot \nabla^*) h|_{\Gamma_{L,R}} = 0, \\ h(x, 1)|_{\Gamma_{top}} &= \phi(x) \quad (\text{Dirichlet}), \\ \mathbf{n} \cdot \mathbf{K} \cdot (D \cdot \nabla^*) h|_{\Gamma_{top}} &= -\mu(h - \phi) \quad (\text{Robin}), \quad x \in [0, 1], \end{aligned} \quad (3.31)$$

where  $\mathbf{n} = (\pm 1, 0)$  are the unit external normal of the lateral surfaces,

$\mathbf{n} = \frac{1}{\sqrt{1 + \phi_{2,x}^2}} (-\phi_{2,x}, 1)$  is the external unit normal of the bottom surface.

Namely,  $\frac{\partial h}{\partial x} \Big|_{left, right} = 0$ . (3.32)  $= 0, \mathbf{n} \cdot \mathbf{K} \cdot (\mathbf{D} \cdot \nabla \phi)$

The consistency condition for the boundary conditions is

$$\int_{\Omega} \mathbf{n} \cdot \mathbf{K} \cdot (\mathbf{D} \cdot \nabla \phi) d\mathbf{x} = \int_{\Omega} \left[ \frac{K_{22}}{\phi - \phi_2} h_x - K_{11} \phi_x (h_x - \frac{\phi_x}{\phi - \phi_2} h_x) \right] d\mathbf{x} = 0. \quad (3.33)$$

We define

$$F(\mathbf{x}, \epsilon) = \left[ \frac{K_{22}}{\phi - \phi_2} h_x - K_{11} \phi_x (h_x - \frac{\phi_x}{\phi - \phi_2} h_x) \right]. \quad (3.34)$$

Analogous to the treatment of the top boundary, we expand  $\phi_2$  in its truncated Fourier Sine series

$$\phi_2 = \phi_2(a) + \frac{\phi_2(b) - \phi_2(a)}{L} (x - a) + \sum_{i=1}^m c_i \sin(i\pi \frac{x-a}{L}). \quad (3.35)$$

We denote

$$\mathbf{h}_t = (\phi(a), \phi(b), b_1, \dots, b_m, \phi_2(a), \phi_2(b), c_1, \dots, c_m). \quad (3.36)$$

The DeepONet is defined by the following:

$$\mathcal{G}(\mathbf{h}_t, \epsilon, \mathbf{x}) = \sum_{k=1}^p \sum_{i=1}^q c_i^k \sigma \left( \sum_{j=1}^{n+m+4} \xi_{ij}^k \mathbf{h}_{t,j} + \xi_{10}^k \epsilon + \theta_i^k \sigma(W_k \cdot \mathbf{x} + \zeta_k) + b_0 \right). \quad (3.37)$$

The loss function is given by

$$\begin{aligned} \mathcal{L}(\theta, \mathbf{h}_t, \epsilon) = & n^{-1} \sum_{i=1}^n [(D \cdot \nabla) \cdot \mathbf{K} \cdot (D \cdot \nabla) \mathcal{G}(\mathbf{h}_t, \epsilon, \mathbf{x}_i^f)]^2 \\ & + \frac{1}{n_t} \sum_{i=1}^{n_t} [\phi(\mathbf{x}_i^t) - \mathcal{G}(\mathbf{h}_t, \epsilon, \mathbf{x}_i^t)]^2 \\ & + \frac{1}{n_l} \sum_{i=1}^{n_l} [G_x(\mathbf{h}_t, \epsilon, \mathbf{x}_i^l)]^2 + \frac{1}{n_r} \sum_{i=1}^{n_r} [G_x(\mathbf{h}_t, \epsilon, \mathbf{x}_i^r)]^2 \\ & + \frac{1}{n_b} \sum_{i=1}^{n_b} [\frac{1}{\phi} G_x \\ & + [\frac{\Delta x}{2} (F(\mathbf{x}_1^l, \epsilon) + 4 \sum_{i \text{ even} \in \{2, \dots, n_t-1\}} F(\mathbf{x}_i^t, \epsilon) \\ & (\mathbf{h}_t, \epsilon, \mathbf{x}_i^b))]^2 \\ & + 2 \sum_{i \text{ odd} \in \{3, \dots, n_t-2\}} F(\mathbf{x}_i^t, \epsilon) + F(\mathbf{x}_{nt}, \epsilon)]^2. \end{aligned} \quad (3.38)$$

where  $\mathbf{x}_i^t$  are the interior points and  $\mathbf{x}_i^k$  are boundary points at the top, left, right, and bottom boundary, respectively. The total loss is given by (3.6) when a set of top and bottom boundaries are given. In practice, the modified loss function is adopted analogous to what we alluded to earlier.

## 4. Results and discussion

We present the numerical results in two scenarios. In the first scenario, we learn the mapping from an arbitrary surface topography to the solution in the basin while the bottom boundary is assumed flat. In the second, we allow the bottom boundary to be an arbitrary shape as well. We have implemented all the methods using PyTorch. For simplicity, we present the results obtained using the spectral representation and the domain

mapping method only in the following. We remark that different surface representations produce the same numerical result.

### 4.1. Results with an arbitrary top boundary

We first present the results obtained using the spectral representation method.

#### 4.1.1. Sampling of the top boundary representation

The top boundary is defined by (3.7) with coefficients or parameters in  $\mathbf{h}_t$ . The sampling of  $\mathbf{h}_t = (\phi(a), \phi(b), b_1, \dots, b_m)$  is carried out as follows

- We sample  $\phi(a)$  uniformly from [0.7, 0.8] and  $\phi(b)$  uniformly from  $[\phi(a) - 0.2, \phi(a) + 0.2]$  to ensure that fluctuations of the boundary function are reasonable geographically.
- For  $h_0(x) = \sum_{j=1}^8 b_j \sin(j\pi x)$  with  $m = 8$ , we sample  $b_1, \dots, b_8$  uniformly from  $[-1, 1]$ , respectively.
- We calculate  $h_{max} = \max_{x \in [0, 1]} \phi(x)$ ,  $h_{min} = \min_{x \in [0, 1]} \phi(x)$ , and  $h_d = h_{max} - h_{min}$ . We sample  $\lambda \in [0, 0.2]$  and then update coefficients  $b_j := \lambda b_j / h_d, j = 1, \dots, 8$ .
- Then, the top boundary surface is well-represented by vector  $\mathbf{h}_t = (\phi(a), \phi(b), b_1, \dots, b_8)$ .
- For illustration purposes, we set  $\epsilon = 0.01$  throughout the paper.

This sampling method makes sure the top boundary fluctuates in  $0.5 < \phi(x)$

1. The larger fluctuation can be done, but it may not be necessary for the realistic geography.

#### 4.1.2. The dataset

The Loss function is defined by summing up all the squared residues of the equation and the boundary conditions as well as a consistency condition that depends on  $\mathbf{h}_t$ . We denote the input of the neural network in the loss function as follows:

$$\mathbf{z} = (\mathbf{h}_t, \mathbf{x}), \quad (4.1)$$

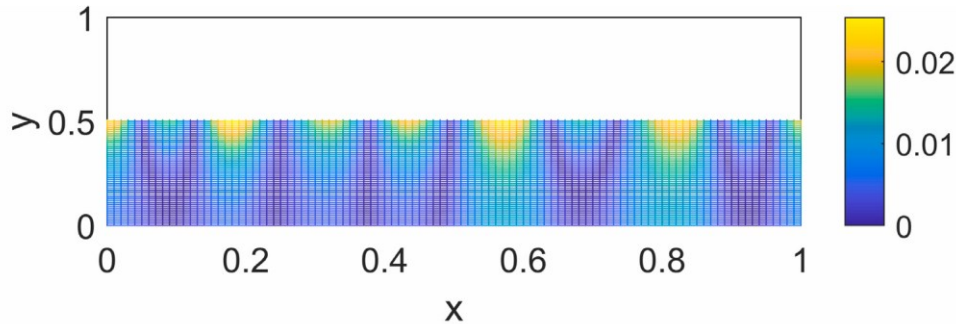
where  $\mathbf{x}$  is a long vector containing randomly chosen points from the boundaries and the interior of the basin underneath the top boundary represented by  $\mathbf{h}_t$  which are chosen after the top boundary is specified. We sample  $l$  number of representing vectors of top boundaries in  $\mathbf{h}_t, l =$

$1, \dots, l$ . For each  $i = 1, \dots, l$ , we have well-defined top boundary  $y = \phi_i(x)$ . For the  $i$ th top boundary, we randomly choose  $L_i$  data points  $\mathbf{x}_j^{t,i}, j = 1, \dots, L_i$  on the left boundary,  $R_i$  points  $\mathbf{x}_j^{r,i}, j = 1, \dots, R_i$  on the right boundary,  $B_i$  points  $\mathbf{x}_j^{b,i}, j = 1, \dots, B_i$  on the bottom, and  $I_i$  points in the interior  $\mathbf{x}_j^i \in [0, 1] \times [0, \phi(\mathbf{x})], j = 1, \dots, I_i, M$  uniform points  $x_{ij}, j = 1, \dots, M$  in  $[0, 1]$ .

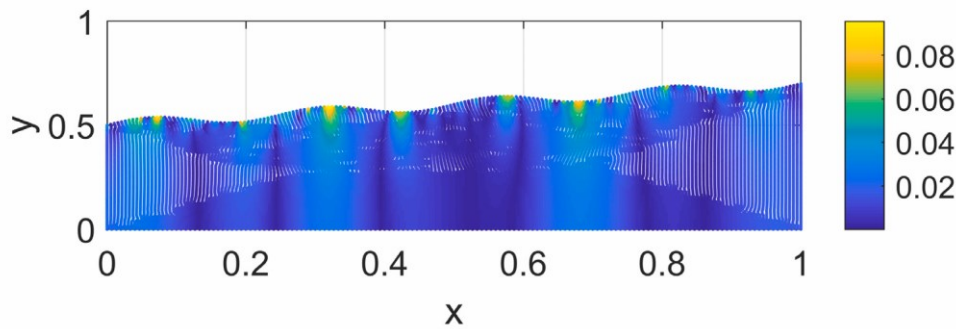
We divide the dataset into the training and test sets by randomly dividing  $\{1, \dots, l\}$  into two subsets  $I_{train}$  and  $I_{test}$ . We generate 140 top boundary topographies using the spectral representation. For 100 boundary topographies, we sample 14000 points randomly, including 10000 interior points and 4000 boundary points. For the rest 40 boundary topographies, we put them in the test set. Finally, we choose  $M = 101$  points uniformly in  $[0, 1]$  to calculate the consistency condition included in the loss function.

In the DeepONet, the width of branch and trunk net is 200, the depth of the branch net is 4, and the depth of the trunk net is 3. We use the Adam algorithm for the first 1000 epoch optimization step with learning rate  $10^{-4}$  and weight decay  $10^{-7}$ . For the remaining epoch, we use the LBFGS algorithm with

learning rate 0.1. The stopping criterion of the LBFGS is for both training and testing processes. The typical loss in each weighted component in the modified loss function at the end of the machine learning process is summarized in Table 4.1. For the parameters in the model, we use characteristic length scales  $L_y=1000$  m and  $L_x=10000$  m, which lead to  $k_{11}=0.01$  and  $k_{22}=1$



(a) The relative error between the solution obtained from the DeepONet and the Toth's solution.  $RMSE = 6.2 \times 10^{-3}$ .



(b) The relative error between the solution obtained from the DeepONet and the numerical one from COMSOL. Here  $RMSE = 2.05 \times 10^{-2}$ .

Fig. 4.1. The top boundary is given by  $y_t = y_0 + c \cdot x + a \sin bx$ , which  $\tan \alpha = c, a/\cos \alpha = \dot{a}, b/\cos \alpha = \dot{b}, y_0 = 0.5, c = 0.02, a = 0.025$ .

Table 4.1

Typical loss values at the end of training and testing, respectively.

| Loss terms   | Training loss         | Test loss             |
|--------------|-----------------------|-----------------------|
| $Loss_{eq}$  | $4.82 \times 10^{-4}$ | $2.35 \times 10^{-3}$ |
| $Loss_l$     | $4.81 \times 10^{-7}$ | $7.57 \times 10^{-6}$ |
| $Loss_r$     | $5.6 \times 10^{-7}$  | $1.71 \times 10^{-5}$ |
| $Loss_b$     | $5.07 \times 10^{-5}$ | $2.18 \times 10^{-5}$ |
| $Loss_t$     | $4.58 \times 10^{-4}$ | $7.77 \times 10^{-3}$ |
| $Loss_{con}$ | $1.69 \times 10^{-5}$ | $3.38 \times 10^{-3}$ |

Table 4.2

Model, neural network and optimization algorithm parameters.

| Parameter                              | Value     |
|----------------------------------------|-----------|
| Width of trunk and branch net          | 200       |
| Depth of branch net                    | 4         |
| Depth of trunk net                     | 3         |
| Weight decay                           | $10^{-7}$ |
| Learning rate for the first 1000 epoch | $10^{-4}$ |
| Learning rate for the remaining epoch  | 0.1       |
| $L_y$                                  | 1000 m    |
| $L_x$                                  | 10000 m   |
| $k_{11}$                               | 0.01      |
| $k_{22}$                               | 1         |

(i.e., aspect ratio  $\varepsilon = 0.01$ ). All model parameters are summarized in Table 4.2.

#### 4.1.3. Benchmark with the Toth water table solution and the numerical solution obtained using COMSOL

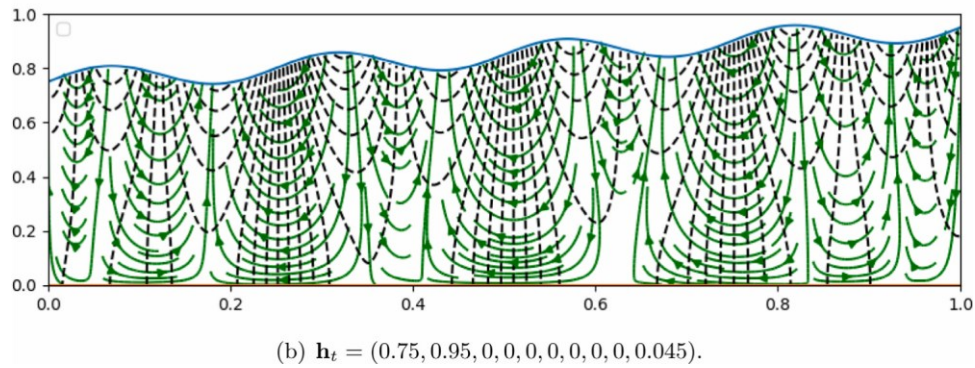
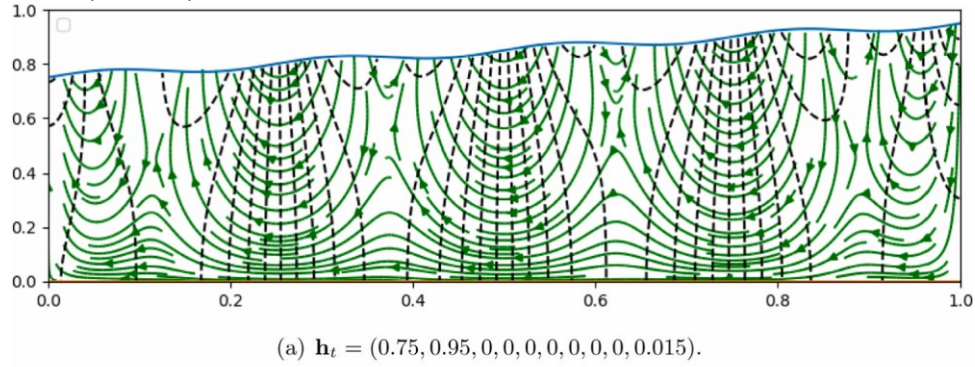
We compare the solution obtained from the DeepONet mapping and the Toth's water table solution and the numerical solution obtained from COMSOL for a given top surface. The results are summarized in Fig. 4.1. The relative

mean square errors (RMSEs) are in the order of  $10^{-2}$ . Given the loss in machine-learning is in the order of  $10^{-3}$ , the Toth's solution is asymptotic, and the numerical solution is approximate, the RMSEs are consistent with the errors one expects from the PIML approach.

#### 4.1.4. Results obtained using the spectral representation

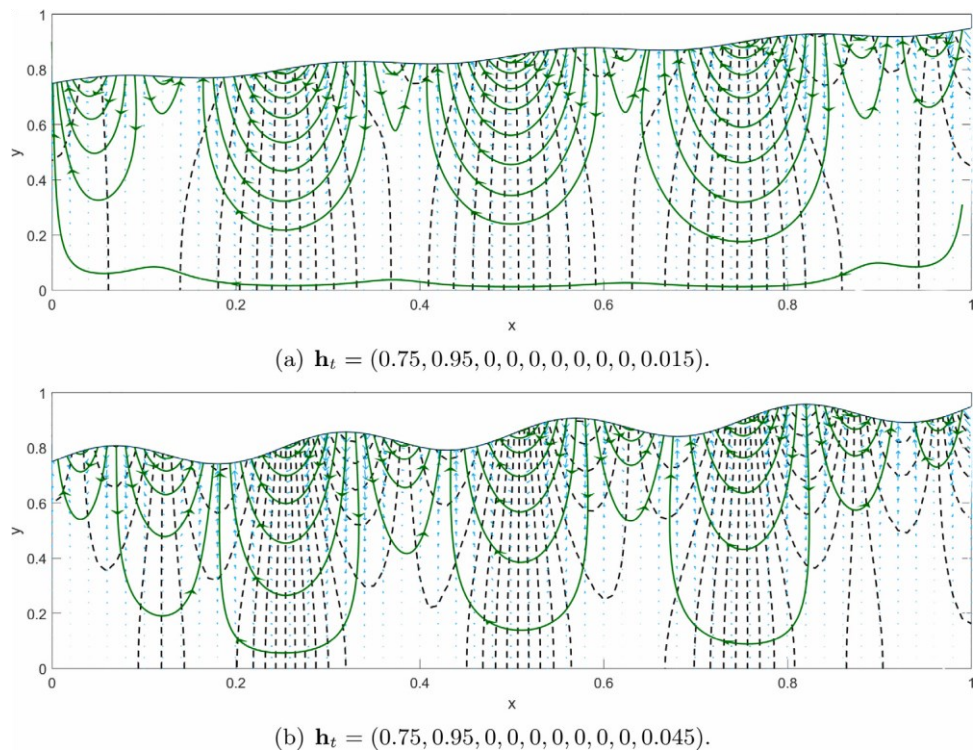
After learning the mapping represented by the DeepONet, we present several representative results obtained using the mapping to show the solution of the steady state groundwater flow equation in the Toth basin. Fig. 4.2 depicts the flow field and the water potential distribution in a Toth basin with sloped top surface topographies of localized variations. There are two factors in the top boundary that impact on flow patterns in the Toth basin: one is the average slope and the other is the localized variation of the surface. A large average slope tends to promote long distance transport of the flow at the bottom of the basin in addition to the compartmentalized or localized circulatory flow patterns near the top surface. When localized variations in the top surface are large, the long distance transport near the bottom tends to be blocked by intruding localized circulations penetrated down from the top. Fig. 4.2 shows a typical steady state flow pattern due to the surface topography with a fixed average slope and varying localized surface variations in space. The two topographical features identified and their influence to steady state flow patterns are shown visibly. To render a better graphical resolution for some long streamlines, we use an image reconstruction method to reconstruct some of the continuous long streamlines that are not well-shown in Fig. 4.2. The results are depicted in Fig. 4.3.

Consistent with Toth's results (Toth, 1963), our solutions also show that a

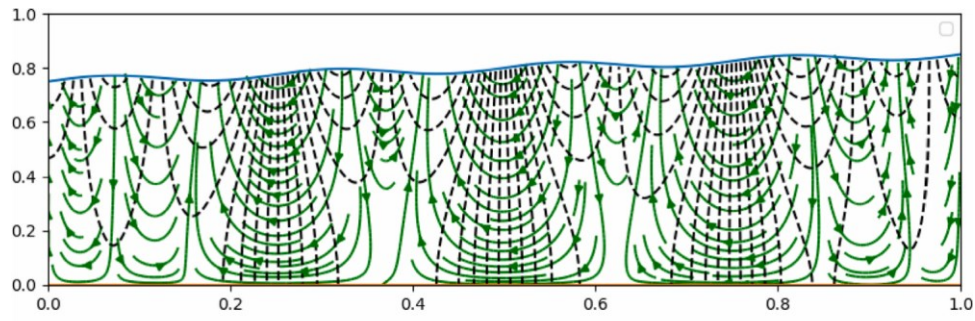


small average slope in the topography and small fluctuation in spatial variations promotes compartmentalized circulations. Fig. 4.4

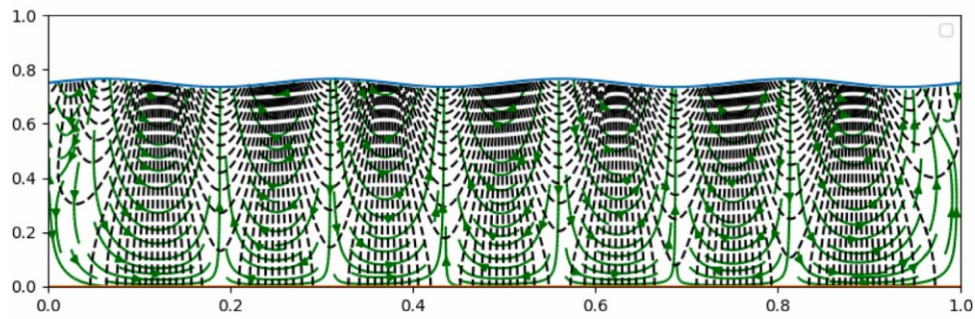
**Fig. 4.2.** Steady flow patterns in streamlines and potential distributions in the Toth basin with a sloped and localized spatial variations at the top boundary represented by a spectral representation. (a). The slow varying surface topography promotes long distance transport at the bottom of the basin due to the slope. (b). When the localized spatial variations in the top are enhanced, the number of compartmentalized circulations increases and flows are blocked to compartmentalized circulations, cutting off the long distance transport.



**Fig. 4.3.** Replot of the steady solutions in Fig. 4.2 with continuous streamlines and potential distributions. These plots reconstruct some continuous long streamlines partially shown in Fig. 4.2.

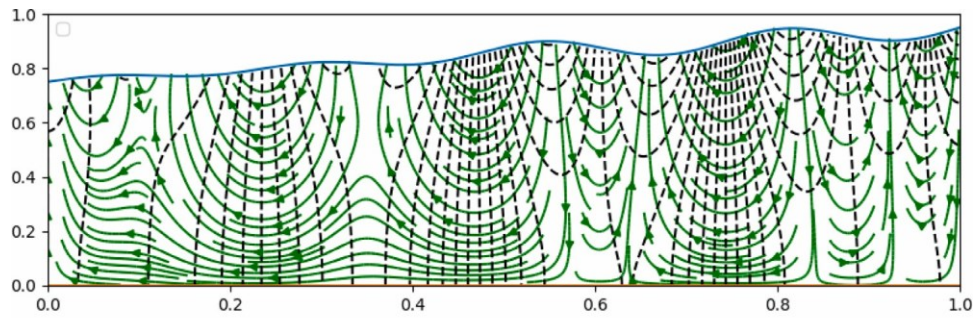


(a)  $\mathbf{h}_t = (0.75, 0.85, 0, 0, 0, 0, 0, 0, 0, 0.015)$ .

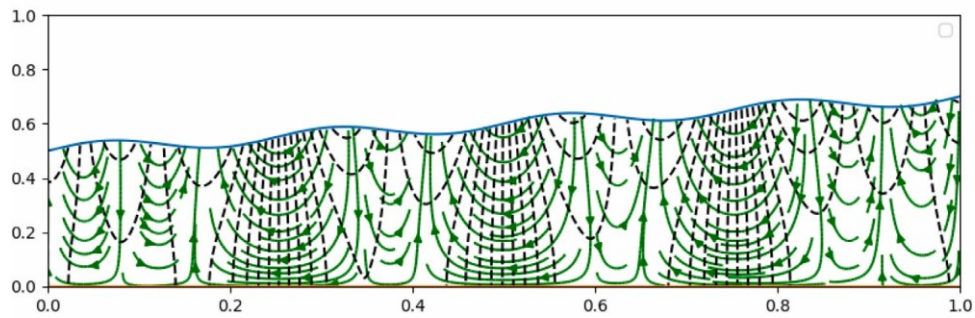


(b)  $\mathbf{h}_t = (0.75, 0.75, 0, 0, 0, 0, 0, 0, 0, 0.015)$ .

**Fig. 4.4.** Steady flow patterns with top surface of small slopes. Small slopes in the surface topography with small surface fluctuations lead to compartmentalized circulations. A top boundary with a zero average slope separates all flows into compartmentalized circulations of nearly equal width.

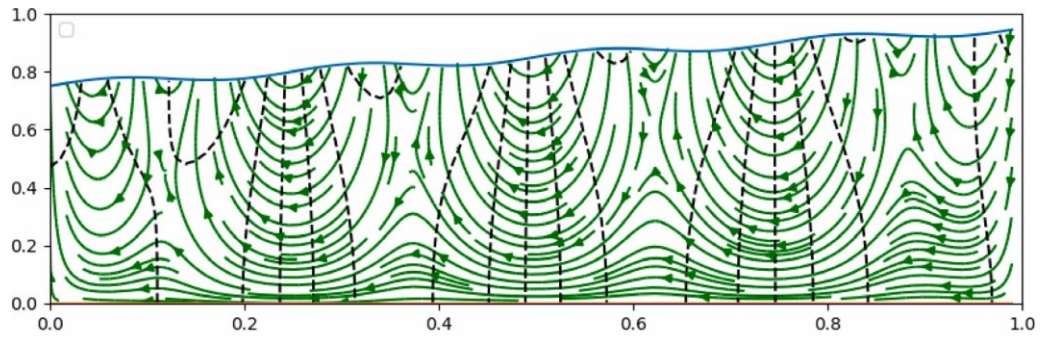


(a)

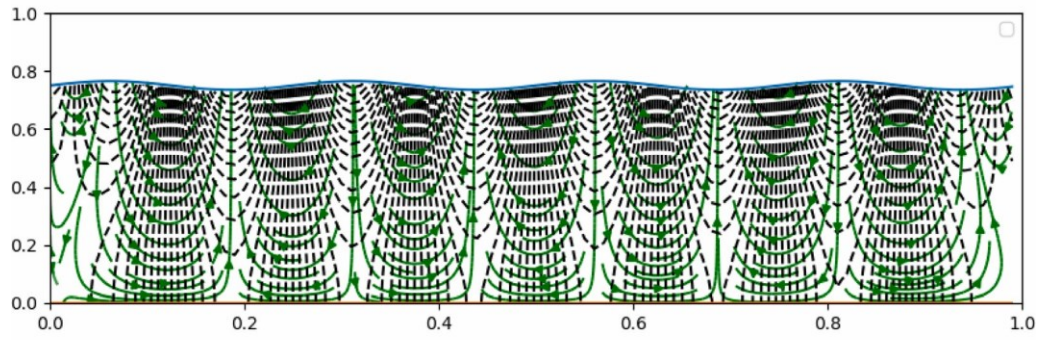


(b)

**Fig. 4.5.** (a). The curvature is smaller on the left than that on the right. Parameter values are  $\mathbf{h}_t = [0.75, 0.95, 3.619 \times 10^{-3}, -1.561 \times 10^{-16}, -4.418 \times 10^{-3}, 4.224 \times 10^{-16}, 5.858 \times 10^{-3}, 1.472 \times 10^{-16}, -1.526 \times 10^{-2}, 2.25 \times 10^{-2}]$ . (b). The curvature enhances on the left in this plot. Parameter values are  $\mathbf{h}_t = [0.5, 0.7, -1.276 \times 10^{-6}, 2.679 \times 10^{-6}, -4.382 \times 10^{-6}, 6.69 \times 10^{-6}, -1.028 \times 10^{-5}, 1.717 \times 10^{-5}, -3.732 \times 10^{-5}, 0.025]$ .

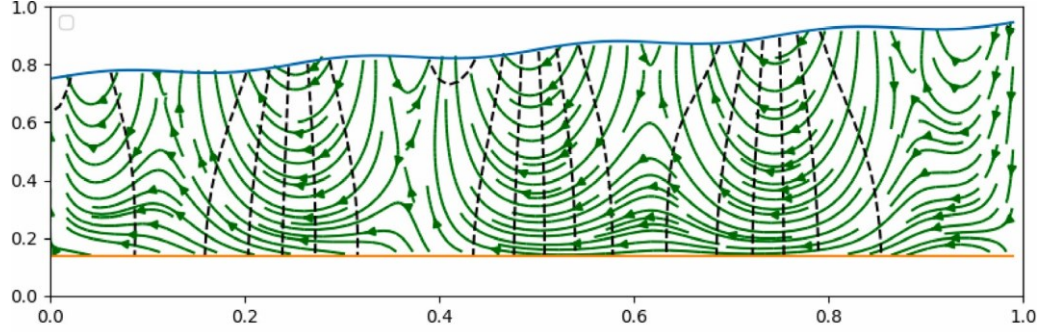


(a)  $\mathbf{h}_t = (0.75, 0.95, 0, 0, 0, 0, 0, 0, 0, 0.015)$ .

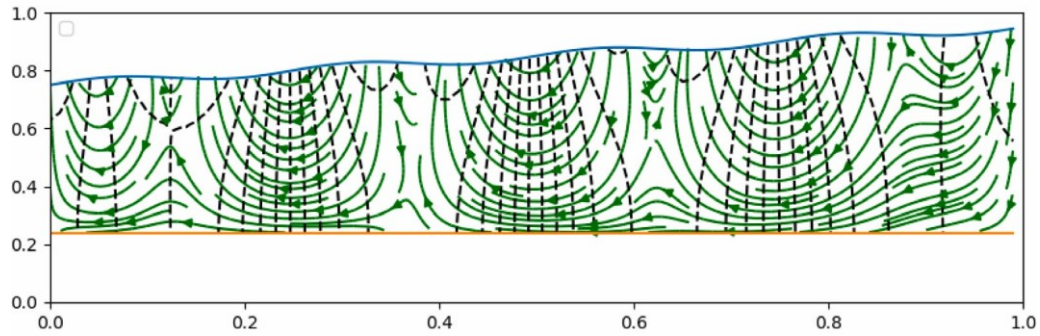


(b)  $\mathbf{h}_t = (0.75, 0.75, 0, 0, 0, 0, 0, 0, 0, 0.015)$ .

**Fig. 4.6.** Steady state solutions obtained using the domain mapping method. (a). This is identical to the one with the same parameters in Fig. 4.2a. (b). This is identical to the one with the same parameters in Figure 4.4b. Thus, the two methods produce the same results.

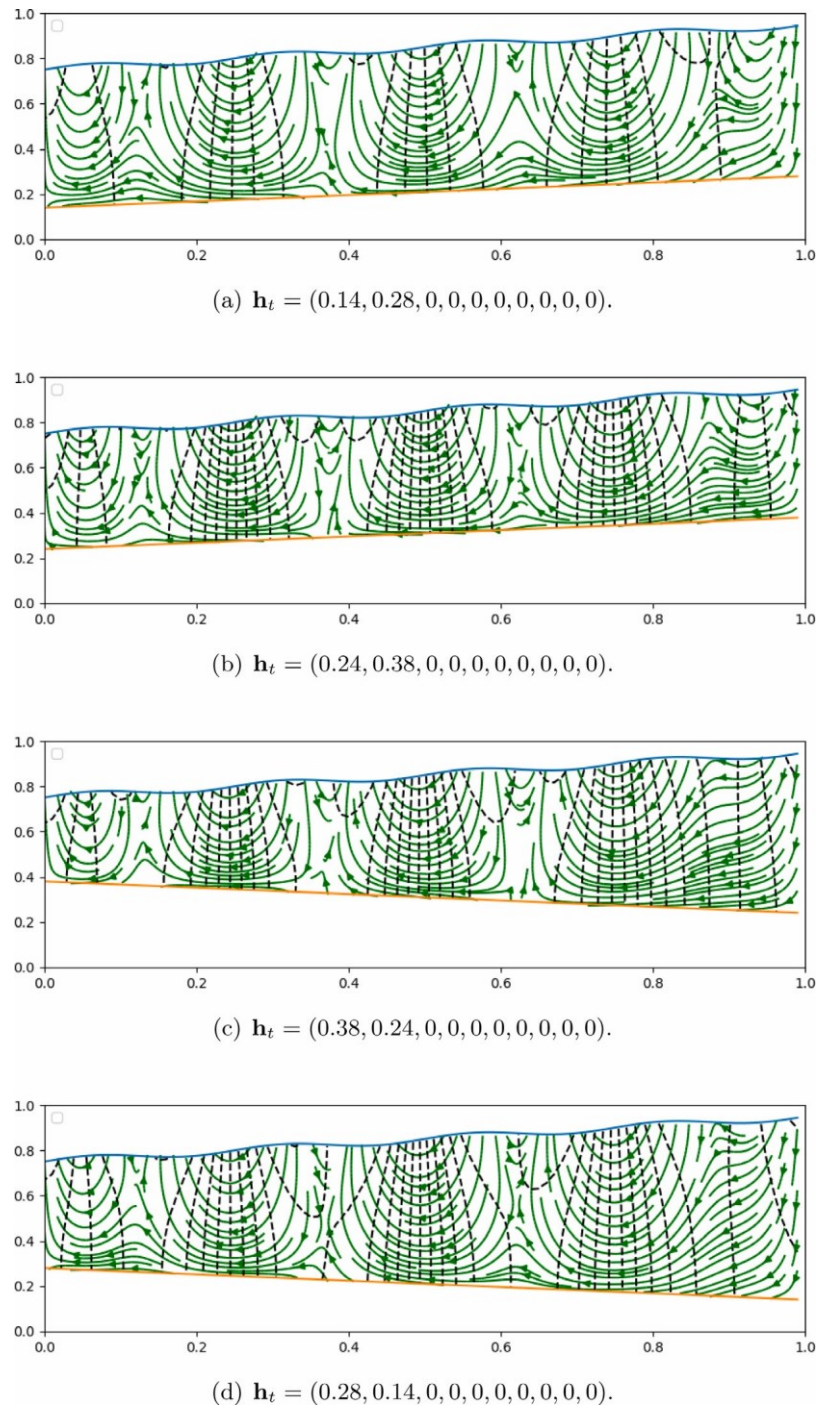


(a)  $\mathbf{h}_t = (0.14, 0.14, 0, 0, 0, 0, 0, 0, 0, 0)$ .



(b)  $\mathbf{h}_t = (0.24, 0.24, 0, 0, 0, 0, 0, 0, 0, 0)$ .

**Fig. 4.7.** The steady state solution in the Toth basin with given top and bottom boundaries represented by spectral representations obtained using the domain mapping approach. (a). The flow pattern is altered compared to Fig. 4.2. When the bottom is lifted, i.e., the basin is shallower, the flows are compressed. (b). When the bottom is lifted further, the flows are compressed further and localization is more prominent. However, the flow pattern does not seem to differ much from (a) qualitatively.



**Fig. 4.8.** The steady state solution in the Toth basin with given top and bottom boundaries represented by spectral representations obtained using the domain mapping approach. (a). The flow has a tendency to flow to the left in long distance due to the increasing slope over there. (b). The flow tends to travel long distance down the bottom boundary to the left analogous to (a). There is no qualitative difference between (a) and (b) where the depth of the basin is different. (c). Flows are more compartmentalized due to the shallow basin. (d). As the depth increases in the basin, the longer range flow is observed near the bottom.

shows two cases of flow fields with small average slopes of the top surface. A top surface with a zero average slope and some spatial variations creates several fully compartmentalized flow patterns correlated with the wave form of the top boundary.

We observe that local curvatures in the top surface affect flow patterns in the Toth basin as well. A larger local curvature tends to create more localized flow patterns while the smaller one promotes more global flow patterns in the bottom of the basin. Fig. 4.5 depicts an example where the magnitude of the curvature of the top surface at the left is smaller than that on the right. As the result, the flow pattern is more localized at the right than that on the left. We

note that it is the overall slope of the surface topography that dominates the overall flow pattern, while the local curvature makes the flow pattern more localized (or circulatory). There apparently exists a competition between the local curvature effect and the overall slope of the top boundary. The flow pattern in the classical Toth water table resembles the flat topographical surface shown in Fig. 4.5b since the Toth's solution is an asymptotic one over a near flat top boundary. The current study indeed extends the asymptotic analyses in Tóth (1963) to a truly nonlinear topography of potentially large spatial variations. With the

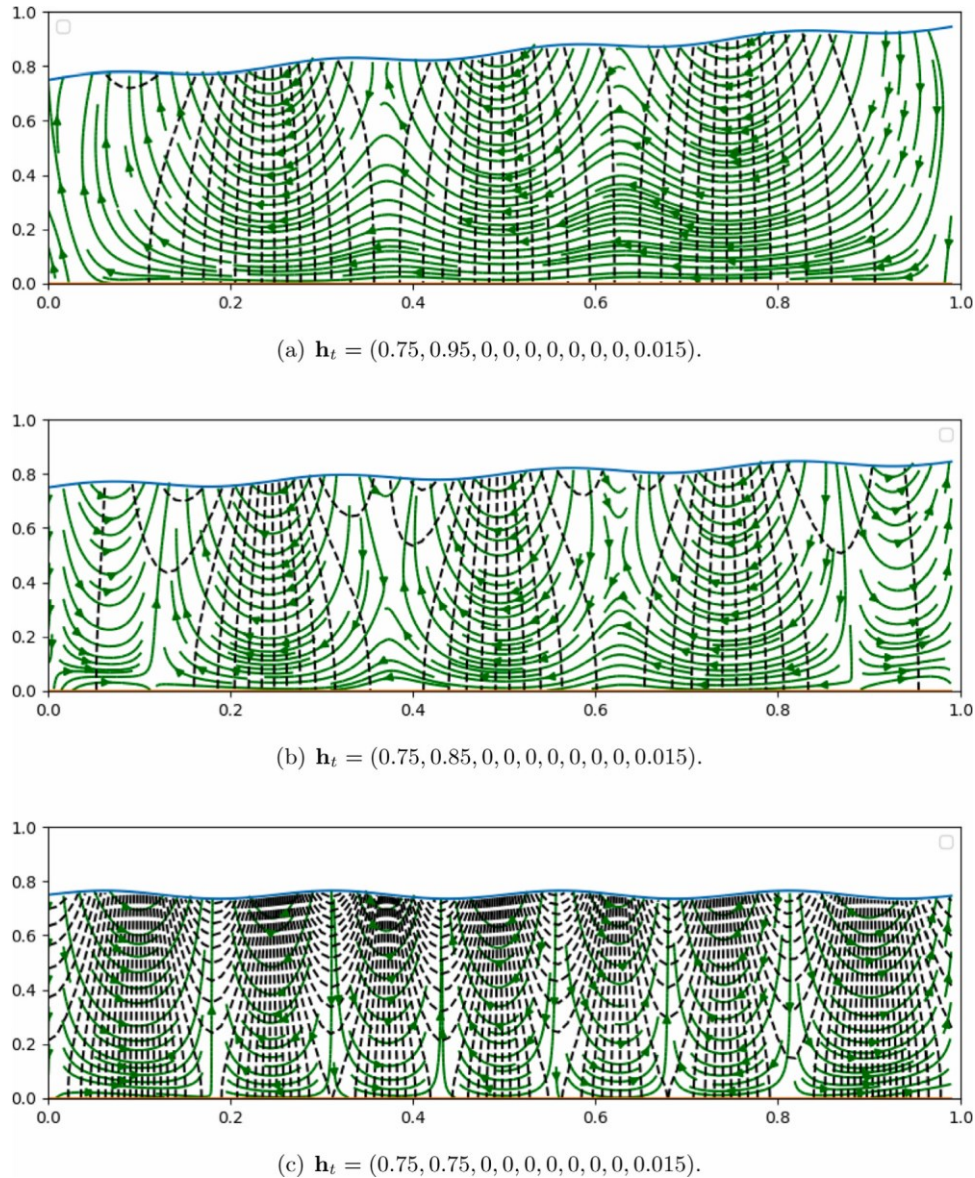


before training to construct the loss function. It could be treated as an input variable for the DeepONet though. But, it would take much longer time to train the neural network in this case. So, we decide not to pursue it in this study. For the Robin boundary condition, we are interested in what changes it brings to the steady flow patterns in comparison to the Dirichlet case.  $\geq$

By examining the numerical results, we observe that when  $\gamma = 10$ , the flow patterns in the three cases with different slopes and spatial variations are pretty much independent of the increase in values of  $\gamma$ . Namely, the patterns

## 5. Remark on inhomogeneous equations and time dependent groundwater flow equations

The method extends readily to the boundary value problem with a source/sink term  $Q$  and the initial-boundary value problem of (2.2) in the framework of PIML. For the inhomogeneous boundary value problem, we modify the loss function by including source term  $Q$  is the penalty terms in the interior equation and the consistency condition, respectively. For the time-dependent initial-boundary value problem, we add one more dimension in



**Fig. 4.10.** Steady flow patterns at three selected slopes of the top boundary with  $\gamma = 10$ . (a). The top boundary of a higher slope produces long-distance flow near the impervious bottom boundary. (b). Compartmentalization becomes more prominent as the slope reduces and in the meantime the number of circulatory flow cell increases. (c). When the average slope of the top surface is zero, long distance flows completely ceases and compartmentalized flow patterns dominate.

in Fig. 4.10 ( $\gamma = 10$ ) are nearly the same as those in Fig. 4.11 when  $\gamma = 10^7$ . While  $\gamma$  becomes smaller however, the flow patterns change quite dramatically, the number of localized flow patterns decreases and long distance flows become more prominent. This is because the penetration length is enhanced in this case so that flow patterns near the bottom are affected directly by the top boundary condition. We expect that the solution is going to be approaching the solution of the Dirichlet problem as  $\gamma \rightarrow \infty$ . The steady states we have calculated do support the observation (see Fig. 4.12).

time and sample randomly in both space and time for the interior and boundary points. The changes that need to be made are in the interior penalty term and the consistency condition term in the loss function. The  $L_2$  norms used in the loss function needs to be expanded to include the average in time. Some additional techniques in training such as time-domain decomposition, etc. need to be considered to accelerate the training process (Lu et al., 2021).

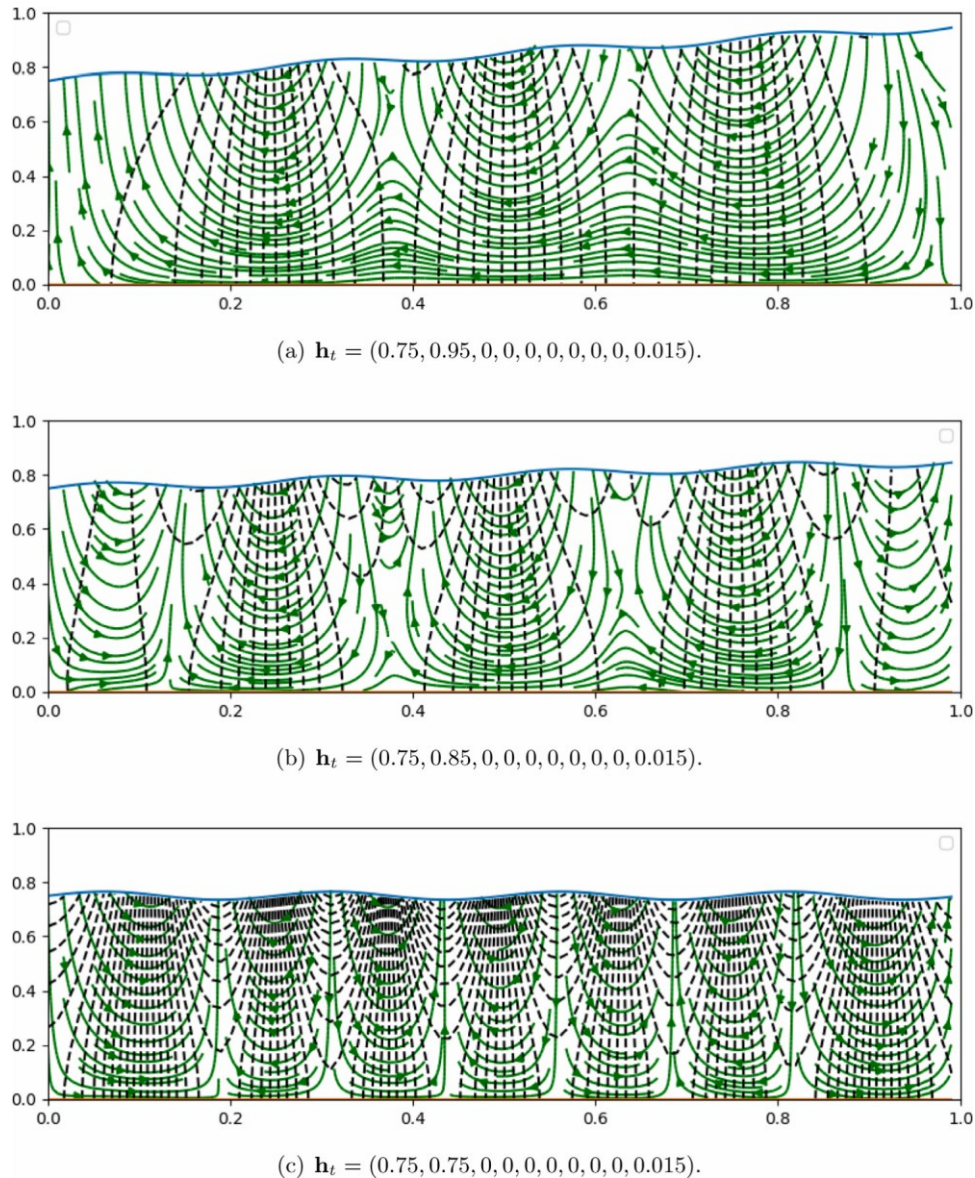
## 6. Conclusion

We have introduced a new approach to obtaining a mapping between surface topography and the solution of the steady state groundwater flow equation in Toth basins with arbitrary surface topographies. This method utilizes PIML with DeepONet and relies on inputting parameters, such as the conductivity tensor and aspect ratio, into the groundwater flow equation. The resulting boundary-to-solution mapping can be learned and repeatedly used to map out the underground steady state flow field for any surface topography and impervious boundary of any shape, providing an alternative means to

used as a well-trained “knowledgeable” boundary-to-solution predictor, demonstrating a neural network representation of an inverse differential operator in hydrological applications, which can be applied in numerous cases.

### CRediT authorship contribution statement

**Jingwei Sun:** Wrote the code, Computed the DeepONet solution, Writing – original draft. **Jun Li:** Worked on the model formulation and solution, Validation, Wrote part of the draft. **Yonghong Hao:** Assisted with the design of the problem and solution, Validation.



**Fig. 4.11.** Steady flow patterns at three selected slopes of the top boundary with  $\gamma = 10^7$ . (a). The top boundary of a higher slope produces long-distance flows near the impervious bottom boundary. (b). Compartmentalization becomes more prominent as the slope reduces and in the meantime the number of circulatory flow cell increases. (c). When the average slope of the top surface is zero, the long distance flow completely ceases and compartmentalized flow patterns dominate. The results are similar to those in Fig. 4.10.

study groundwater flow phenomena in complex geophysical systems. This method can estimate groundwater flow patterns in new Toth basins with similar geophysical parameters, and new models can be efficiently machine-learned through transfer learning with a predetermined architecture even in locations where geophysical parameters differ. With additional computational efforts, the model parameters can also be treated as an input to the DeepONet, expanding the mapping to a wider range of groundwater flow equations with varying parameter values. As a result, the DeepONet can be

**Cuiting Qi:** Computed the COMSOL solution. **Chunmei Ma:** Assisted with problem design and student, Supervision. **Huazhi Sun:** Assisted with problem design and student, Supervision. **Negash Begashaw:** Assisted with problem design and student, Supervision. **Gurcan Comert:** Assisted with problem design and student, Supervision. **Yi Sun:** Assisted with problem design and student, Supervision. **Qi Wang:** Designed the problem, Worked on all aspects of the project, Writing – original draft.

## Declaration of competing interest

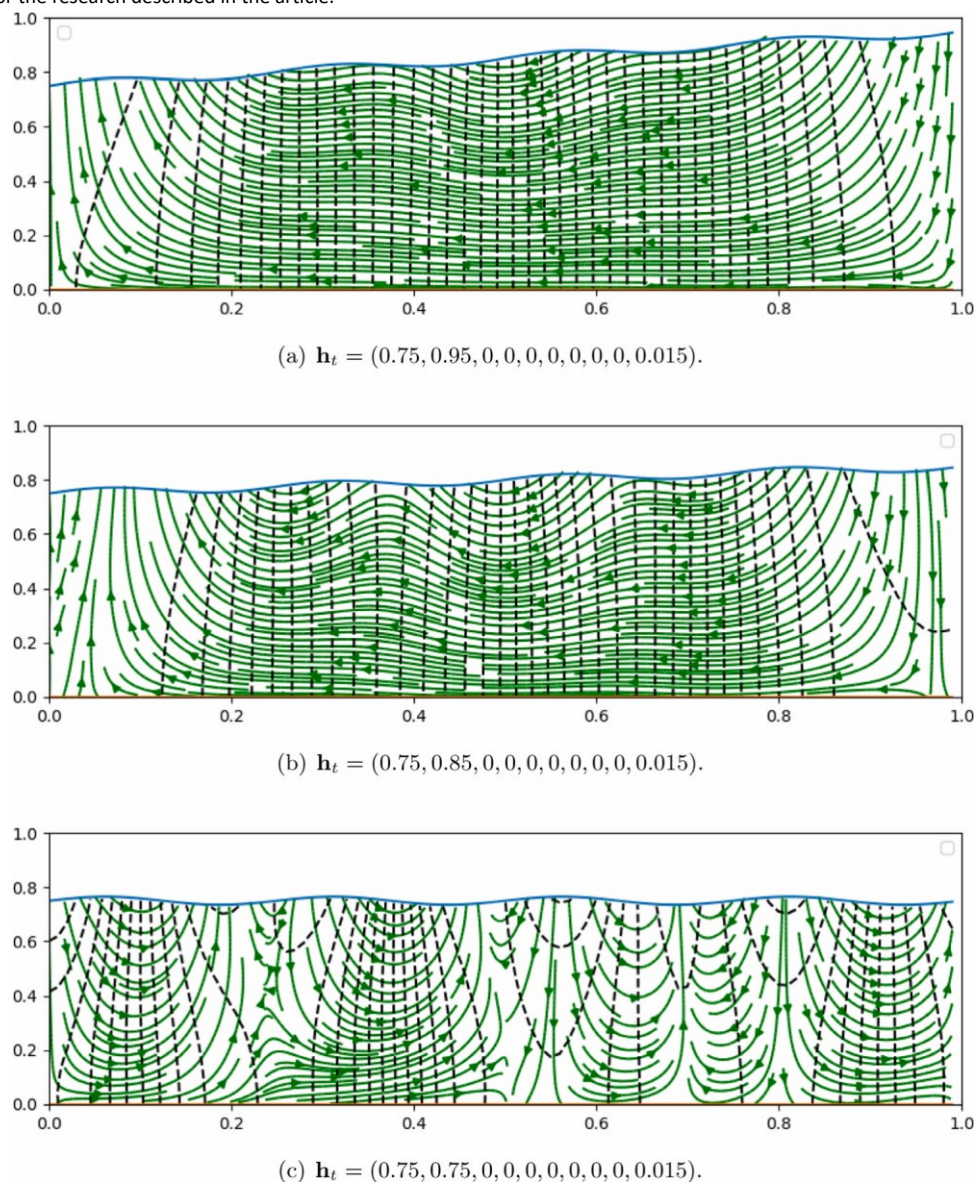
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

No data was used for the research described in the article.

## References

- An, Ran, Jiang, Xiao Wei, Wang, Jun Zhi, Wan, Li, Wang, Xu Sheng, Li, Hailong, 2014. A theoretical analysis of basin-scale groundwater temperature distribution. *Hydrogeol. J.* 23 (2), 397–404.
- Butcher, J.C., 2016. *Numerical Methods for Ordinary Differential Equations*, third ed. Wiley, Chichester, West Sussex, United Kingdom.
- Cao, Shuhao, 2021. Choose a Transformer: Fourier or Galerkin. p. 17, arXiv 2105. 14995.
- Cardenas, M. Bayani, Jiang, Xiao Wei, 2010. Groundwater flow, transport, and residence times through topography-driven basins with exponentially decreasing permeability and porosity. *Water Resour. Res.* 46 (11), 3–23.
- Chen, Tianping, Chen, Hong, 1995. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems. *IEEE*



**Fig. 4.12.** Steady flow patterns at three selected slopes of the top boundary with  $\gamma = 0.2$ . (a). The top boundary of a higher slope produces long-distance flows near the impervious bottom boundary. (b). Compartmentalization appears near the top locally. (c). When the average slope of the top surface is zero, the long distance flow across the entire bottom of the basin ceases. However, the compartmentalized flow cells are larger than in the case where  $\gamma = 10$ .

## Acknowledgments

Jun Li's work is partially supported by grants from The Science & Technology Development Fund of Tianjin Education Commission for Higher Education 2020KJ005, and the National Natural Science Foundation of China 11971247. Negash Begashaw, Gurcan Comert, Yi Sun, and Qi Wang's work is partially supported by National Science Foundation award DMS-1954532.

- Trans. Neural Netw. 6 (4), 911–917.
- Evans, Lawrence C., 2010. *Partial Differential Equations*, second ed. AMS.
- Guibas, John, Mardani, Morteza, Li, Zongyi, Tao, Andrew, Aanandkumar, Anima, Catanzaro, Bryan, 2022. Adaptive fourier neural operators: Efficient token mixers for transformers. p. 15, arXiv, 2111.13587.
- Haberman, Richard, 2013. *Applied Partial Differential Equations: With Fourier Series and Boundary Value Problems*, fifth ed. PEARSON, Boston.
- Jiang, X.-W., Wang, X.-S., Wan, L., Ge, S., 2011. An analytical study on stagnation points in nested flow systems in basins with depth-decaying hydraulic conductivity. *Water Resour. Res.* 47, W01512.

- Karniadakis, George Em, Kevrekidis, Ioannis G., Lu, Lu, Perdikaris, Paris, Wang, Sifan, Yang, Liu, 2021. Physics-informed machine learning. *Nat. Rev. Phys.* 3 (6), 422–440.
- Kissas, Georgios, Seidman, Jacob, Guilhoto, Leonardo Ferreira, Preciado, Victor M., Pappas, George J., Perdikaris, Paris, 2022. Learning operators with coupled attention. p. 39, arXiv, 2201.01032.
- Leshno, Moshe, Lin, Vladimir Ya, Pinkus, Allan, Schocken, Shimon, 1993. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. *Neural Netw.* 6 (6), 861–867.
- Li, Zongyi, Kovachki, Nikola, Azizzadenesheli, Kamyar, Liu, Burigede, Bhattacharya, Kaushik, Stuart, Andrew, Anandkumar, Anima, 2021a. Fourier Neural Operator for parametric partial differential equations. arXiv, 2010.08895.
- Li, Yakun, Yu, Wenkai, Zhao, Jia, Wang, Qi, 2021b. Second order linear decoupled energy dissipation rate preserving schemes for the Cahn-Hilliard-extended-Darcy model. *J. Comput. Phys.* 444 (4), 110561.
- Lu, Lu, Jin, Pengzhan, Karniadakis, George Em, 2019. DeepONet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators. arXiv, 1910.03193.
- Lu, Lu, Jin, Pengzhan, Pang, Guofei, Zhang, Zhongqiang, Karniadakis, George Em, 2021. Learning nonlinear operators via deepONet based on the universal approximation theorem of operators. *Nat. Mach. Intell.* 3, 218–229.
- Nield, Donald A., Bejan, Adrian, 2006. *Convection in Porous Media*, third ed. Springer, New York.
- Niu, Hong, Liang, Xing, Ni, Sheng Nan, Wen, Zhang, 2015. Analytical study of unsteady nested groundwater flow systems. *Math. Probl. Eng.*
- Pang, Guofei, Lu, Lu, Karniadakis, George Em, 2019. fPINNs: Fractional physics-informed neural networks. *SIAM J. Sci. Comput.* 41 (4), A2603–A2626.
- Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* 378, 686–707.
- Sukumar, N., Srivastava, Ankit, 2022. Exact imposition of boundary conditions with distance functions in physics-informed deep neural networks. *Comput. Methods Appl. Mech. Engrg.* 389, 114333.
- Tóth, J., 1962. A theory of groundwater motion in small drainage basins in central Alberta, Canada. *J. Geophys. Res.* 67 (11), 4375–4388.
- Tóth, J., 1963. A theoretical analysis of groundwater flow in small drainage basins. *J. Geophys. Res.* 68 (16), 4795–4812.
- Toth, J., 1970. A conceptual model of the groundwater regime and the hydrogeologic environment. *J. Hydrol.* 10 (2), 164–176.
- Wang, Xu Sheng, Wan, L., Jiang, Xiao Wei, Li, H., Zhou, Y., Wang, J., Ji, X., 2017. Identifying three-dimensional nested groundwater flow systems in a tothian basin. *Adv. Water Resour.* 108, 139–156.