Approximation of the Boundary-to-Solution Operator for the Groundwater Transport Equation in a Toth Basin

Jingwei Sun[1], Jun Li[1], Yonghong Hao[1], Huazhi Sun[1], Chunmei Ma[1], Yi sun[2], Gurcan Comert[3], Negash Negashaw[3], and Qi Wang[2]

[1]Tianjin Normal University
[2]University of South Carolina
[3]Benedict College

November 22, 2022

Abstract

We develop a deep learning approach to learn the boundary-to-solution operator, i.e., to establish the boundary to steady solution mapping, in the Toth basin of arbitrary top and bottom topographies and two types of prescribed boundary conditions.
The machine-learned mapping is represented by a DeepOnet, which takes the geometrical data and boundary conditions as the inputs and produces the steady state solution as the output. In this approach, we approximate the top and bottom boundaries by either truncated Fourier series or piecewise linear representations. The DeepOnet maps directly the finite dimensional representations of the boundaries to the steady state solution of the ground water transport equation in the Toth basin. We present two different implementations of the DeepOnet: 1) the Toth basin is embedded in a rectangular computational domain, and 2) the Toth basin with arbitrary top and bottom boundaries is mapped into a rectangular computational domain via a nonlinear transformation. We 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. Both implementations yield the same results, showcasing a new deep learning approach to study ground water transport phenomena.

Approximation of the Boundary-to-Solution Operator for the Groundwater Transport Equation in a Toth Basin

Jingwei Sun[†], Jun Li[††], Yonghong Hao[◇], Chunmei Ma[†], Huazhi Sun[†], Negash Begashaw[◇◇], Gurcan Comet[◇◇], Yi Sun[‡], and Qi Wang[‡]

† School of Computer Science, Tianjin Normal University, Tianjin, China,100084;

†† School of Mathematical Science, Tianjin Normal University, Tianjin, China, 300387;

◇ Center for Ground Water Research, Tianjin Normal University, Tianjin, China,100084;

◇◇ Mathematics, Computer Science, Physics and Engineering Department, Benedict College, Columbia, SC 20204, USA;

‡* Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA; Email: qwang@math.sc.edu.

June 23, 2022

Abstract

We develop a deep learning approach to learn the boundary-to-solution operator, i.e., to establish the boundary to steady solution mapping, in the Toth basin of arbitrary top and bottom topographies and two types of prescribed boundary conditions. The machine-learned mapping is represented by a DeepOnet, which takes the geometrical data and boundary conditions as the inputs and produces the steady state solution as the output. In this approach, we approximate the top and bottom boundaries by either truncated Fourier series or piecewise linear representations. The DeepOnet maps directly the finite dimensional representations of the boundaries to the steady state solution of the ground water transport equation in the Toth basin. We present two different implementations of the DeepOnet: 1) the Toth basin is embedded in a rectangular computational domain, and 2) the Toth basin with arbitrary top and bottom boundaries is mapped into a rectangular computational domain via a nonlinear transformation. We 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. Both implementations yield the same results, showcasing a new deep learning approach to study ground water transport phenomena.

Keywords: Machine learning, Poisson equation, Boundary-to-solution mapping, Toth basin, DeepOnet, Ground water transport.

1 Introduction

The Toth groundwater analysis was a seminal theoretical attempt to relate surface topography 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 [1, 2]. It involved solving an elliptic boundary value problem for a given surface topography not far from a horizontally flat surface with the associated Dirichlet boundary condition on a rectangular domain approximately. For a general non-rectangular 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 transport in the basin domain approximately. The traditional numerical solver for the solution, while the surface topography is given, only produces one solution for each given topographical surface. When one studies another Toth basin with a different surface topography, the solution would have to be recalculated. One thus wonders if there exists a "solver", that once obtained can map the surface topography directly to the solution in the interior so that the "solver" can produce the solution from a given surface topography.

In the past, one study had been conducted to optimize the Toth theory by refining the coefficient of permeability or the viscosity of the fluid in porus media, or to examine the influence of temperature [3]; and another focused on investigating the influence of depth and systematic heterogeneity [4]. A couple of other studies focused on generalizing the Toth theory to other settings, like the more realistic three-dimensional space [5] and unsteady situations [6] etc. But these works didn't examine the impact of the top surface topography 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 appropriate boundary conditions, and propose a novel idea to establish a mapping from surface topographies (top alone or top+ bottom) to the solution of the governing system of equations for ground water transport in the Toth basin directly using a deep learning method. This is an analogue of a solution formula for an initial-boundary value problem for partial differential equations (PDEs) in terms of deep learning, where the solution of the initial-boundary value problem is expressed as a function of the domain, the boundary conditions and the nonhomogeneous forcing term. This approach will produce a solution "solver" that maps the prescribed initial and boundary conditions as well as the forcing term to the solution directly, suitable for any geophysical basins that share the same transport property such as the mobility coefficient in the transport equation.

The recent advancement in deep learning with neural networks makes the development of such a desired mapping plausible [7–13]. Given that a neural network is a mapping composed of compound functions with specific structures, the mapping should be able to be established should we propose the proper architecture of the deep neural network in principle [14,15]. We note that the governing steady state equation for ground water transport in porus media is a Poisson equation. Given a boundary and physically consistent boundary conditions, a solution can be represented by an integral containing the Green's function [16]. The integral with the Green's function yields the mapping from the initial-boundary conditions and the source term to the solution theoretically. Motivated by this connection between the domain boundary, boundary conditions and the source of the steady state 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 [17]. It is therefore an appropriate tool for us to establish the desired mapping.

Specifically, we will address the following questions in this study using deep learning:

- 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 Ω?
- What is the specific effect of both the top and bottom boundary conditions to the solution

in the Toth basin? 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 a gravity acceleration, and $y = \phi$ defines the top boundary; and (ii) the Robin boundary condition: $\frac{\partial h}{\partial \mathbf{n}}(\mathbf{x}) + \gamma h(\mathbf{x}) = g\phi(x)$, where γ 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.

- What's the direct mapping from the surface topography and basin geometry to the steady state Darcy velocity field in the basin?

We will address these issues altogether by solving the PDE boundary value problem with respect to two distinct boundary conditions using the DeepONet [8, 18]. The presentation is given for steady states without a source term. However, the method extends readily to transient situations and also the transport phenomenon with a source. We note that for a completely new Toth basin, an analogous boundary-to-solution mapping to describe ground water flows in porous media can be developed using transfer learning, which could be more efficiently done.

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. Without loss of generality, we will detail the latter two implementations.

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 three distinctive neural network models for mapping the surface topography together with the prescribed boundary condition to the solution of the governing equation for ground water transport in steady states.

2.1 Model formulation

We first formulate the ground water transport model in a general time-dependent setting. We consider transport of ground water in a given domain Ω with piecewise smooth boundary $\partial\Omega$, 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

$$h(\mathbf{x}, t) = gy + \int_{p_0}^{p} \frac{1}{\rho} dp, \tag{2.1}$$

where g is the gravity acceleration, y is the height of the water basin measured from the bottom impervious layer, $\rho(p)$ is the density of water assumed a function of pressure p, p_0 is the atmospheric pressure at the top surface 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 y is chosen as the lowest point along the bottom surface when it is not flat. The transport equation of $h(\mathbf{x}, t)$ is given by the following continuity equation:

$$S\frac{\partial h}{\partial t} = \nabla \cdot \mathbf{v} + Q, \tag{2.2}$$

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

$$\nabla h = \nabla(gy) + \frac{1}{\rho}\nabla p. \tag{2.3}$$

The constitutive equation between water potential h and Darcy velocity $\mathbf{v}$ is given by the Darcy's law [19]:

$$\mathbf{v} = K \cdot \nabla h, \tag{2.4}$$

where K is the mobility 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

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

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 Ω, this partial differential equation is accompanied by a set of boundary conditions over domain boundary $\Gamma = \partial\Omega$. We consider the following 2D domain with boundary conditions given below (see Figure 2.4),

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

where $\mathbf{n} = \frac{1}{\sqrt{1+\phi_x^2}}(-\phi_x, 1)$ is the unit external normal to the boundary, $\Gamma_{b,l,r}$ are the boundaries at the bottom, left and right side of domain Ω, respectively, and equation $y = \phi(x)$ defines the top boundary. The lateral boundaries are assumed vertical line segments in domain Ω while the top and bottom ones can be of arbitrary shapes. 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 [2]. When the bottom boundary is flat and top boundary inclined with a small slope, Toth calculated his well-known Toth water table in [20] using an approximate analytical method based on an asymptotic analysis on a rectangular domain.

Given any boundary conditions along $\partial\Omega$, we need to check their consistence with the governing equation in Ω [21]. We integrate equation (2.5) over Ω to obtain

$$\int_\Omega [S\frac{\partial h}{\partial t} - (\nabla \cdot (K \cdot \nabla h) + Q)]d\mathbf{x} = \int_\Omega [S\frac{\partial h}{\partial t} - Q]d\mathbf{x} - \int_{\partial\Omega} \mathbf{n} \cdot (K \cdot \nabla h)ds = 0. \tag{2.7}$$

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

$$\int_\Omega [S\frac{\partial h}{\partial t} - Q]d\mathbf{x} - \int_{\Gamma_t} \mathbf{n} \cdot (K \cdot \nabla h)ds = 0, \tag{2.8}$$

where Γ_t is the top boundary of Ω. In steady states and without the source term, in particular,

the consistent condition further reduces to

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

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

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

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

Assuming the boundary 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:

$$\|\hat{h} - h\|_\Omega \leq C_1 \|\nabla \cdot K \cdot \nabla \hat{h}\|_\Omega + C_2 \|\mathbf{n} \cdot K \cdot \nabla \hat{h}\|_{\Gamma - \Gamma_t} + C_3 \|\hat{h} - g\phi\|_{\Gamma_t}, \tag{2.11}$$

where the norms are some proper norms defined in their respective spaces and $C_i, i = 1, 2, 3$ are positive constants [22]. Then,

$$h = arg \min_{\hat{h}} \|\hat{h} - h\|_\Omega = arg \min_{\hat{y}} [C_1 \|\nabla \cdot K \cdot \nabla \hat{h}\|_\Omega + C_2 \|\mathbf{n} \cdot K \cdot \nabla \hat{h}\|_{\Gamma - \Gamma_t} + C_3 \|\hat{h} - g\phi\|_{\Gamma_t}]. \tag{2.12}$$

Thus, we use the righthand side to define the loss function in this case.

$$Loss = C_1 \|\nabla \cdot K \cdot \nabla \hat{h}\|_\Omega + C_2 \|\mathbf{n} \cdot K \cdot \nabla \hat{h}\|_{\Gamma - \Gamma_t} + C_3 \|\hat{h} - g\phi|_{\Gamma_t}. \tag{2.13}$$

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 the physic-informed machine learning (PIML) formulation [7]. 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 [23].

2.3 Robin boundary value problem

A more physical boundary condition in steady states at the top boundary perhaps should be

$$\mathbf{n} \cdot K \cdot \nabla h = -\gamma (h - g\phi(x)), \tag{2.14}$$

where γ 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 $\gamma = 0$, (2.14) reduces to the impervious Neumann boundary condition; whereas it reduces to the Dirichlet one if $\gamma \to \infty$.

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

$$\|h - \hat{h}\| \leq C_1 \|\nabla \cdot K \cdot \nabla \hat{h}\|_\Omega + C_2 \|\mathbf{n} \cdot K \cdot \nabla \hat{h}\|_{\Gamma - \Gamma_t} + C_3 \|\mathbf{n} \cdot K \cdot \nabla h + \gamma (h - g\phi(x))\|_{\Gamma_t}, \tag{2.15}$$

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

$$Loss = C_1\|\nabla \cdot K \cdot \nabla \hat{h}\|_{\Omega} + C_2\|\mathbf{n} \cdot K \cdot \nabla \hat{h}\|_{\Gamma - \Gamma_t} + C_3\|\mathbf{n} \cdot K \cdot \nabla h + \gamma(h - g\phi(x))\|_{\Gamma_t}. \quad (2.16)$$

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 (2.14) yields the following consistency condition:

$$\int_{\Gamma_t} \gamma(h - \phi) ds = 0. \quad (2.17)$$

(2.9) and (2.17) are two constraints for the solution to satisfy the Dirichlet and the Robin boundary condition, respectively.

2.4 Nondimensionalizaton

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 by

$$\tilde{x} = \frac{x}{L_x}, \tilde{y} = \frac{y}{L_y}, \tilde{t} = \frac{t}{T}, \tilde{h} = \frac{h}{h_0}, \tilde{\phi} = \frac{\phi}{L_y}, \quad (2.18)$$

where h_0 is a characteristic water potential. We denote the characteristic storage rate by S_0. The dimensionless model parameters are defined by

$$\tilde{S} = \frac{S}{S_0}, \tilde{K} = \frac{T}{L_y^2 S_0}\mathbf{A} \cdot K \cdot \mathbf{A}, \tilde{Q} = \frac{TQ}{S_0 h_0}, \quad (2.19)$$

where

$$\mathbf{A} = \begin{pmatrix} \epsilon & 0 \\ 0 & 1 \end{pmatrix}, \quad (2.20)$$

$\epsilon = \frac{L_y}{L_x}$ is the aspect ratio of the basin.

The transport equation in the dimensionless form is given by

$$\tilde{S}\frac{\partial \tilde{h}}{\partial \tilde{t}} = \nabla \cdot \tilde{K} \cdot \nabla \tilde{h} + \tilde{Q}. \quad (2.21)$$

The top Dirichlet boundary condition is given by

$$\tilde{h} = \frac{gL_y}{h_0}\tilde{\phi}. \quad (2.22)$$

We choose

$$h_0 = gL_y, \quad (2.23)$$

and drop the ˜ from the dimensionless equations to obtain the dimensionless equation and boundary

conditions as follows:

$$\begin{aligned} &S\frac{\partial h}{\partial t} = \nabla \cdot K \cdot \nabla h + Q, \mathbf{x} \in \Omega, \\ &h = \phi(x), \mathbf{x} \in \Gamma_t, \quad \mathbf{n} \cdot K \cdot \nabla h = 0, \mathbf{x} \in \partial\Omega - \Gamma_t. \end{aligned} \tag{2.24}$$

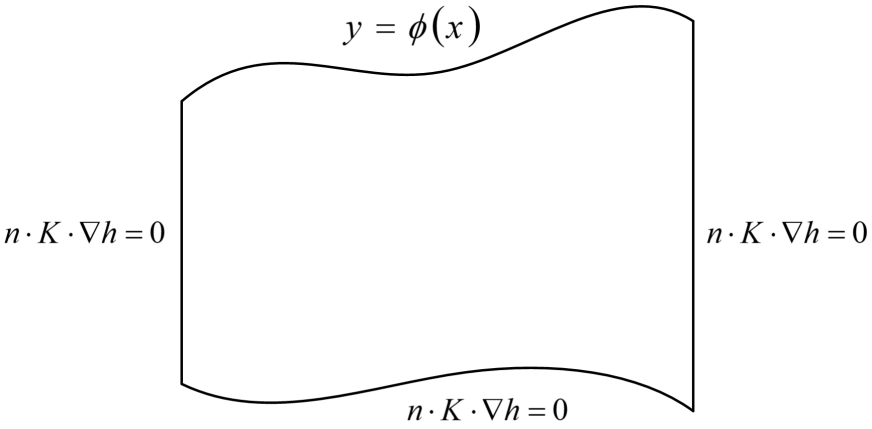
The consistent condition (2.9) retains.

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

$$\mathbf{n} \cdot K\nabla h = -\tilde{\gamma}(h - \phi), \tag{2.25}$$

where $\tilde{\gamma} = \gamma L_x$. We drop the tilde over γ for brevity.

In this paper, we consider $K = Diag(K_{11}, K_{22})$ as a diagonal mobility matrix, $x \in [0, 1]$, and $0 < \phi_2(\mathbf{x}) < \phi(\mathbf{x}) \leq 1$ in the dimensionless equation and boundary conditions, where $y = \phi_2$ is the bottom boundary.



(a)

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

Next, we present three DeepOnet implementations for the mapping from specified boundary and the boundary conditions (2.6) to the steady state water potential solution and thereby the flow field via the Darcy's law in Ω [17].

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 steady solution in the domain. We adopt the physics-informed machine learning approach and use the DeepOnet as the neural network to represent the mapping. The DeepOnet is specifically designed for this purpose [17]. For the Toth basin, we are interested in a domain with flat lateral, arbitrary top and bottom boundaries. Firstly, we consider the bottom boundary is also flat and impervious so that the zero Neumann boundary conditions are valid at all lateral and bottom boundaries. The top boundary is defined by $y = \phi(x)$ with aspect ratio ϵ. We construct a mapping from the top surface boundary

condition of the water potential to the steady state solution of the governing equation in the Toth basin. Owing to the special boundary condition in this particular problem, the boundary representation and the boundary condition coincide. So, we only need to learn a mapping from top boundary $y = \phi(x)$ to the solution of the steady state governing equation in Ω. We present three different approaches to accomplishing this goal using two distinct representations of the top boundary $y = \phi(x)$, respectively.

3.1 Piecewise polynomial interpolation of the top boundary

Firstly, we represent the top boundary, $y = \phi(x)$, using n discrete points $\mathbf{x}_i = (x_i, \phi(x_i)), i = 1, \cdots, n$, uniformly distributed in the x-coordinate, where $x_i = (i-1)\Delta x, \Delta x = \frac{1}{n-1}$. We denote the approximate solution of this mixed boundary value problem in the interior of Ω by a DeepONet $G(\mathbf{h}_t, \epsilon, \mathbf{x})$ as follows

$$G(\mathbf{h}_t, \epsilon, \mathbf{x}) = \sum_{k=1}^{p} \sum_{i=1}^{q} c_i^k \sigma(\sum_{j=1}^{n} \xi_{ij}^k \phi(\mathbf{x}_j) + \xi_{i0}^k \epsilon + \theta_i^k) \sigma(W_k \cdot \mathbf{x} + \zeta_k) + b_0, \tag{3.1}$$

where c_i^k, ξ_{ij}^k, W_k are weights and θ_i^k, ζ_k, b_0 are biases of the neural network, $\mathbf{h}_t = (\phi(x_1), \cdots, \phi(x_n)) \in \mathbb{R}^n$ denotes the uniformly distributed, y-coordinates of the interpolating points at the top boundary, and n, p, q are positive integers. We use the DeepOnet to learn the mapping from $\mathbf{h}_t, \epsilon$ to the solution. We choose n_l, n_r, n_b, n_i points at the left, right, bottom boundary, and the interior randomly.

We denote $n_t = n$ (odd) as the number of points at the top boundary. The loss function of the machine learning model is given by (2.13) with the L_2 norms. We evaluate the integral norms using the Monte Carlo sampling. For the randomly chosen points along the boundary and in the interior, $\{\mathbf{x}_j^i, i = l, r, b, \mathbf{x}_j\}$, and a well-defined uniform division of $[0, 1]$, $\{x_j, j = 1, \cdots, n_t\}, \mathbf{x}_j^t = (x_j, \phi(x_j))$, the specific expression of the loss function is given by

$$L(\theta, \mathbf{h}_t, \epsilon) = \tfrac{1}{n} \textstyle\sum_{i=1}^{n} (\phi(\mathbf{x}_i) - G(\mathbf{h}_t, \epsilon, \mathbf{x}_i^t))^2 + \tfrac{1}{n_i} \sum_{j=1}^{n_i} (\nabla \cdot K \cdot \nabla G(\mathbf{h}_t, \epsilon, \mathbf{x}_j))^2 +$$

$$\tfrac{1}{n_l} \textstyle\sum_{i=1}^{n_l} (\frac{\partial}{\partial x} G(\mathbf{h}_t, \epsilon, \mathbf{x}_i^l))^2 + \tfrac{1}{n_r} \sum_{i=1}^{n_r} (\frac{\partial}{\partial x} G(\mathbf{h}_t, \epsilon, \mathbf{x}_i^r))^2 + \tfrac{1}{n_b} \sum_{i=1}^{n_b} (\frac{\partial}{\partial y} G(\mathbf{h}_t, \epsilon, \mathbf{x}_i^b))^2 \tag{3.2}$$

$$+ [\tfrac{\Delta x}{3} \big(F(\mathbf{h}_t, \epsilon, \mathbf{x}_1) + 4 \textstyle\sum_{i \ even \in \{2, \cdots, n-1\}} F(\mathbf{h}_t, \epsilon, \mathbf{x}_i^t) + 2 \sum_{i \ odd \in \{3, \cdots, n-2\}} F(\mathbf{h}_t, \epsilon, \mathbf{x}_i^t) + F(\mathbf{h}_t, \epsilon, \mathbf{x}_n^t) \big)]^2,$$

where $\{\mathbf{x}_j\}_{j=1}^{n_i}$ are the interior points and $\{\mathbf{x}_i^k\}_{i=1}^{n_k}$ are boundary points, $k = t, l, r, b$ represent the top, left, right, and bottom boundary, respectively, and the boundary mass flux is given by

$$F(\mathbf{h}_t, \epsilon, \mathbf{x}) = \mathbf{n}^{(t)} \cdot K \cdot (\frac{\partial}{\partial x} G(\mathbf{h}_t, \epsilon, \mathbf{x}), \frac{\partial}{\partial y} G(\mathbf{h}_t, \epsilon, \mathbf{x}))^T = -K_{11} \phi_x(\mathbf{x}) G_x(\mathbf{h}_t, \epsilon, \mathbf{x}) + K_{22} G_y(\mathbf{h}_t, \epsilon, \mathbf{x}) \tag{3.3}$$

The Simpson's quadrature formula is employed [24] to ensure the integral is accurate up to the fourth order in Δx. 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 $\mathbf{h}_t$ 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 re-balancing the weights. The loss function is then modified into

$$L(\theta,\mathbf{h}_t,\epsilon)=\frac{\lambda_1}{n}\sum_{i=1}^{n}(\phi(\mathbf{x}_i)-G(\mathbf{h}_t,\epsilon,\mathbf{x}_i^t))^2+\frac{\lambda_2}{n_i}\sum_{j=1}^{n_i}(\nabla\cdot K\cdot\nabla G(\mathbf{h}_t,\epsilon,\mathbf{x}_j))^2+$$

$$\frac{\lambda_3}{n_l}\sum_{i=1}^{n_l}(\frac{\partial}{\partial x}G(\mathbf{h}_t,\epsilon,\mathbf{x}_i^l))^2+\frac{\lambda_4}{n_r}\sum_{i=1}^{n_r}(\frac{\partial}{\partial x}G(\mathbf{h}_t,\epsilon,\mathbf{x}_i^r))^2+\frac{\lambda_5}{n_b}\sum_{i=1}^{n_b}(10\times\frac{\partial}{\partial y}G(\mathbf{h}_t,\epsilon,\mathbf{x}_i^b))^2 \tag{3.4}$$

$$+\lambda_6[\frac{\Delta x}{3}\big(F(\mathbf{h}_t,\epsilon,\mathbf{x}_1)+4\sum_{i\ even\in\{2,\cdots,n-1\}}F(\mathbf{h}_t,\epsilon,\mathbf{x}_i^t)+2\sum_{i\ odd\in\{3,\cdots,n-2\}}F(\mathbf{h}_t,\epsilon,\mathbf{x}_i^t)+F(\mathbf{h}_t,\epsilon,\mathbf{x}_n^t)\big)]^2,$$

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

$$loss_{top}=\frac{1}{n}\sum_{i=1}^{n}(\phi(\mathbf{x}_i)-G(\mathbf{h}_t,\epsilon,\mathbf{x}_i^t))^2,\quad loss_{eq}=\frac{1}{n_i}\sum_{j=1}^{n_i}(\nabla\cdot K\cdot\nabla G(\mathbf{h}_t,\epsilon,\mathbf{x}_j))^2,$$
$$loss_{left}=\frac{1}{n_l}\sum_{i=1}^{n_l}(G_x(\mathbf{h}_t,\epsilon,\mathbf{x}_i^l))^2,\quad loss_{right}=\frac{1}{n_r}\sum_{i=1}^{n_r}(G_x(\mathbf{h}_t,\epsilon,\mathbf{x}_i^r))^2,$$
$$loss_{bottom}=\frac{1}{n_b}\sum_{i=1}^{n_b}(G_y(\mathbf{h}_t,\epsilon,\mathbf{x}_i^b))^2,$$
$$loss_{con}=[\frac{\Delta x}{3}\big(F(\mathbf{h}_t,\epsilon,\mathbf{x}_1)+4\sum_{i\ even\in\{2,\cdots,n-1\}}F(\mathbf{h}_t,\epsilon,\mathbf{x}_i^t)+2\sum_{i\ odd\in\{3,\cdots,n-2\}}F(\mathbf{h}_t,\epsilon,\mathbf{x}_i^t)+F(\mathbf{h}_t,\epsilon,\mathbf{x}_n^t)\big)]^2, \tag{3.5}$$
$$\overline{loss^{i-1}}=(loss_{top}^{i-1}+loss_{eq}^{i-1}+loss_{right}^{i-1}+loss_{left}^{i-1}+loss_{bottom}^{i-1}+loss_{con}^{i-1})/6,$$
$$\lambda_1^i=loss_{top}^{i-1}/\ \overline{loss^{i-1}},\quad \lambda_2^i=loss_{eq}^{i-1}/\ \overline{loss^{i-1}},\quad \lambda_3^i=loss_{right}^{i-1}/\ \overline{loss^{i-1}},$$
$$\lambda_4^i=loss_{left}^{i-1}/\ \overline{loss^{i-1}},\quad \lambda_5^i=loss_{bottom}^{i-1}/\ \overline{loss^{i-1}},\quad \lambda_6^i=loss_{con}^{i-1}/\ \overline{loss^{i-1}}$$

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

$$L(\theta)=\frac{1}{m}\sum_{l=1}^{L}\sum_{i=1}^{m}L(\theta,\mathbf{h}_t^{(i)},\epsilon^{(l)}). \tag{3.6}$$

We remark that the numbers of randomly selected interior and boundary points at each given $\mathbf{h}_t$ and $\epsilon^{(l)}$ are not the same so that $L(\theta,\mathbf{h}_t^{(i)},\epsilon^{(l)})$ can have different number of terms in the square sums. Note that the 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's 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 the continuous top boundary $\phi(x)$ using a truncated Sine Fourier series together with a linear interpolation function as follows [16]:

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

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

$$b_j = \frac{2}{L}\int_0^L [\phi(x) - \phi(0) - \frac{\phi(L)-\phi(0)}{L}x]\sin(\frac{j\pi}{L}x)dx. \tag{3.8}$$

We then construct a mapping from the top boundary given by (3.7) to the solution in the Toth basin. We represent the top boundary using $m+2$ discrete values $\mathbf{h}_t = (\phi(0), \phi(L), b_1, \cdots, b_m) \in \mathbb{R}^{m+2}$, consisting of the Sine Fourier coefficients and the two end point values. Given the boundary condition at the top boundary $\phi(\mathbf{x})$, we want to learn a mapping from $\mathbf{h}_t$ to the solution of the steady state governing equation in Ω.

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

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

where c_i^k, ξ_{ij}^k, W_k are weights and θ_i^k, ζ_k, b_0 are 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 Ω. We divide $[0,1]$ uniformly into $n-1$ intervals, separated by $x_i = (i-1)\Delta x, \Delta x = \frac{1}{n-1}, i = 1, \cdots, n$. We use the DeepOnet to learn the mapping from $\mathbf{h}_t$ to solution $h(\mathbf{x}, t)$ in Ω. The cost function in the model for each given top boundary, the randomly chosen points along the boundary and in the interior, $\{\mathbf{x}_j^i, i = l, r, b, \mathbf{x}_j\}$, a well-defined uniformed division of $[0,1]$, $\{x_j, j = 1, \cdots, n_t\}$, that defines the top boundary points $\mathbf{x}_j^t = (x_j, \phi(x_j))$, and aspect ratio ϵ is then defined by

$$\begin{aligned} L(\theta, \mathbf{h}_t, \epsilon) &= \tfrac{1}{n_i}\textstyle\sum_{j=1}^{n_i}(\nabla\cdot K\cdot\nabla G(\mathbf{h}_t, \epsilon, \mathbf{x}_j))^2 + \tfrac{1}{n_t}\sum_{j=1}^{n_t}[\phi(\mathbf{x}_j^t) - G(\mathbf{h}_t, \epsilon, \mathbf{x}_j^t)]^2 \\ &+ \tfrac{1}{n_l}\textstyle\sum_{i=1}^{n_l}(G_x(\mathbf{h}_t, \epsilon, \mathbf{x}_i^l))^2 + \tfrac{1}{n_r}\sum_{i=1}^{n_r}(G_x(\mathbf{h}_t, \epsilon, \mathbf{x}_i^r))^2 + \tfrac{1}{n_b}\sum_{i=1}^{n_b}(G_y(\mathbf{h}_t, \epsilon, \mathbf{x}_i^b))^2 \\ &+ [\tfrac{\Delta x}{3}\big(F(\mathbf{x}_1, \epsilon) + 4\textstyle\sum_{i\ even\in\{2,\cdots,n-1\}} F(\mathbf{x}_i^t, \epsilon) + 2\sum_{i\ odd\in\{3,\cdots,n-2\}} F(\mathbf{x}_i^t, \epsilon) + F(\mathbf{x}_n^t, \epsilon)\big)]^2, \end{aligned} \tag{3.10}$$

where $\mathbf{x}_j$ are the interior points and $\mathbf{x}_j^k, k = t, l, r, b$ are boundary points at the top, left, right, and bottom boundary, respectively. In the practical implementation, we once again re-balance the "local loss" as alluded to earlier.

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

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

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

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

Or equivalently,

$$\begin{aligned} \nabla &= (\tfrac{\partial}{\partial x}, \tfrac{\partial}{\partial y}) = (\tfrac{\partial}{\partial x} - \tfrac{y}{\phi^2}\phi_x \tfrac{\partial}{\partial z}, \tfrac{1}{\phi}\tfrac{\partial}{\partial z}) \\ &= \begin{pmatrix} 1 & -\frac{y}{\phi^2}\phi_x \\ 0 & \frac{1}{\phi(x)} \end{pmatrix} \cdot \nabla^*. \end{aligned} \tag{3.13}$$

We denote

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

The Laplace equation is rewritten into

$$(D \cdot \nabla^*) \cdot K \cdot (D \cdot \nabla^*) h = 0. \tag{3.15}$$

The boundary conditions of h is given by

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

In this study, we limit ourselves to

$$K = Diag(K_{11}, K_{22}). \tag{3.17}$$

Then,

$$\begin{aligned} &\mathbf{n} \cdot K \cdot (D \cdot \nabla^*) h|_{\Gamma_{l,r}} = K_{11} \tfrac{\partial h}{\partial x}|_{\Gamma_{l,r}} = 0, \\ &\mathbf{n} \cdot K \cdot (D \cdot \nabla^*) h|_{\Gamma_b} = \tfrac{K_{22}}{\phi} \tfrac{\partial h}{\partial z}|_{\Gamma_b} = 0. \end{aligned} \tag{3.18}$$

These imply

$$\frac{\partial h}{\partial x}|_{\Gamma_{l,r}} = 0, \quad \frac{\partial h}{\partial z}|_{\Gamma_b} = 0, \tag{3.19}$$

where the bottom boundary is assumed flat.

The steady state governing equation without a source is given by

$$\begin{aligned} &(D \cdot \nabla^*) \cdot K \cdot (D \cdot \nabla^*) h \\ &= K_{11}[\tfrac{\partial^2 h}{\partial x^2} - \tfrac{\partial}{\partial x}(\tfrac{z\phi_x}{\phi}\tfrac{\partial h}{\partial z}) - \tfrac{z\phi_x}{\phi}\tfrac{\partial^2 h}{\partial x \partial z} + \tfrac{z\phi_x}{\phi}\tfrac{\partial}{\partial z}(\tfrac{z\phi_x}{\phi}\tfrac{\partial h}{\partial z})] + \tfrac{K_{22}}{\phi}\tfrac{\partial}{\partial z}(\tfrac{1}{\phi}\tfrac{\partial h}{\partial z}) = 0. \end{aligned} \tag{3.20}$$

The consistent condition becomes

$$\int_{z=1}[\mathbf{n}^{(t)}\cdot K\cdot(D\cdot\nabla^*)h]\,dx=\int_{z=1}[K_{22}\frac{h_z}{\phi}-K_{11}\phi_x(h_x-\frac{\phi_x}{\phi}h_z)]dx=0, \tag{3.21}$$

where $\mathbf{n}^{(t)}=(-\phi_x,1)$. We then construct a mapping from the dimensionless (and also parameterized) equation to the solution in the Toth basin. We represent the top boundary using $m+2$ discrete values $\mathbf{h}_t=(\phi(a),\phi(b),b_1,\cdots,b_m)$ from the truncated Sine Fourier series approximation. Given the boundary condition at the top boundary $z=1$, we want to learn a mapping from $\mathbf{h}_t$ to the solution of the steady state governing equation in Ω.

We denote the solution in the interior of Ω by the DeepONet $G(\mathbf{h}_t,\mathbf{x})$ defined in (3.9). The loss function for each given top boundary $\mathbf{h}_t$, the randomly chosen points along the boundary and in the interior, $\{\mathbf{x}_j^i,i=l,r,b,\mathbf{x}_k\}$, and a well-defined uniform division of $[0,1]$, $\{x_j,\}$, that defines the top boundary points $\mathbf{x}_j^t=(x_j,1)$, is then defined by

$$\begin{aligned}L(\theta,\mathbf{h}_t,\epsilon)&=\tfrac{1}{n_i}\textstyle\sum_{j=1}^{n_i}[(D\cdot\nabla^*)\cdot K\cdot(D\cdot\nabla^*)G(\mathbf{h}_t,\epsilon,\mathbf{x}_j^t)]^2+\tfrac{1}{n_t}\sum_{j=1}^{n_t}[\phi(\mathbf{x}_j^t)-G(\mathbf{h}_t,\epsilon,\mathbf{x}_j^t)]^2\\&+\tfrac{1}{n_l}\textstyle\sum_{i=1}^{n_l}[K_{11}G_x(\mathbf{h}_t,\epsilon,\mathbf{x}_i^l)]^2+\tfrac{1}{n_r}\sum_{i=1}^{n_r}[K_{11}G_x(\mathbf{h}_t,\epsilon,\mathbf{x}_i^r)]^2+\tfrac{1}{n_b}\sum_{i=1}^{n_b}[\mathbf{n}\cdot K\cdot(D\cdot\nabla^*)G(\mathbf{h}_t,\epsilon,\mathbf{x}_i^b)]^2\\&+[\tfrac{\Delta x}{3}(F(\mathbf{h}_t,\epsilon,\mathbf{x}_1^t)+4\textstyle\sum_{i\ even\in\{2,\cdots,n_t-1\}}F(\mathbf{h}_t,\epsilon,\mathbf{x}_i^t)+2\sum_{i\ odd\in\{3,\cdots,n_t-2\}}F(\mathbf{h}_t,\epsilon,\mathbf{x}_i^t)+F(\mathbf{h}_t,\epsilon,\mathbf{x}_{n_t}))]^2,\end{aligned} \tag{3.22}$$

where $\mathbf{x}_j$ are the interior points and $\mathbf{x}_j^k,k=t,l,r,b$ are boundary points at the top, left, right, and bottom boundary, respectively,

$$F(\mathbf{x},\epsilon)=K_{22}\frac{G_z}{\phi}-K_{11}\phi_x(G_x-\frac{\phi_x}{\phi}G_z). \tag{3.23}$$

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

$$\begin{aligned}L(\theta,\mathbf{h}_t,\epsilon)&=\tfrac{\lambda_1}{n_t}\textstyle\sum_{j=1}^{n_t}[\phi(\mathbf{x}_j^t)-G(\mathbf{h}_t,\epsilon,\mathbf{x}_j^t)]^2+\tfrac{\lambda_2}{n_i}\sum_{j=1}^{n_i}[(D\cdot\nabla^*)\cdot K\cdot(D\cdot\nabla^*)G(\mathbf{h}_t,\epsilon,\mathbf{x}_j)]^2\\&+\tfrac{\lambda_3}{n_l}\textstyle\sum_{i=1}^{n_l}[G_x(\mathbf{h}_t,\epsilon,\mathbf{x}_i^l)]^2+\tfrac{\lambda_4}{n_r}\sum_{i=1}^{n_r}[G_x(\mathbf{h}_t,\epsilon,\mathbf{x}_i^r)]^2+\tfrac{\lambda_5}{n_b}\sum_{i=1}^{n_b}[10\times\mathbf{n}\cdot K\cdot(D\cdot\nabla^*)G(\mathbf{h}_t,\epsilon,\mathbf{x}_i^b)]^2\\&+\lambda_6[\tfrac{\Delta x}{3}(F(\mathbf{x}_1^t,\epsilon)+4\textstyle\sum_{i\ even\in\{2,\cdots,n_t-1\}}F(\mathbf{x}_i^t,\epsilon)+2\sum_{i\ odd\in\{3,\cdots,n_t-2\}}F(\mathbf{x}_i^t,\epsilon)+F(\mathbf{x}_{n_t},\epsilon))]^2,\end{aligned} \tag{3.24}$$

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

$$
\begin{aligned}
&loss_{top} = \tfrac{1}{n_t}\textstyle\sum_{i=1}^{n_t}(\phi(\mathbf{x}_i^t) - G(\mathbf{h}_t, \mathbf{x}_i^t))^2, \quad loss_{eq} = \tfrac{1}{n_i}\textstyle\sum_{j=1}^{n_i}((D\cdot\nabla^*)\cdot K\cdot(D\cdot\nabla^*)G(\mathbf{h}_t, \mathbf{x}_j))^2,\\
&loss_{left} = \tfrac{1}{n_l}\textstyle\sum_{i=1}^{n_l}(G_x(\mathbf{h}_t, \mathbf{x}_i^l))^2, \quad loss_{right} = \tfrac{1}{n_r}\textstyle\sum_{i=1}^{n_r}(G_x(\mathbf{h}_t, \mathbf{x}_i^r))^2,\\
&loss_{bottom} = \tfrac{1}{n_b}\textstyle\sum_{i=1}^{n_b}(\tfrac{K_{22}}{\phi}G_z(\mathbf{h}_t, \mathbf{x}_i^b))^2,\\
&loss_{con} = [\tfrac{\Delta x}{3}(F(\mathbf{x}_1^t) + 4\textstyle\sum_{i\ even\in\{2,\cdots,n_t-1\}}F(\mathbf{x}_i^t) + 2\textstyle\sum_{i\ odd\in\{3,\cdots,n_t-2\}}F(\mathbf{x}_i^t) + F(\mathbf{x}_{n_t}))]^2 \qquad (3.25)\\
&\overline{loss^{i-1}} = (loss_{top}^{i-1} + loss_{eq}^{i-1} + loss_{right}^{i-1} + loss_{left}^{i-1} + loss_{bottom}^{i-1} + loss_{con}^{i-1})/6,\\
&\lambda_1^i = loss_{top}^{i-1}/\ \overline{loss^{i-1}}, \quad \lambda_2^i = loss_{eq}^{i-1}/\ \overline{loss^{i-1}}, \quad \lambda_3^i = loss_{right}^{i-1}/\ \overline{loss^{i-1}},\\
&\lambda_4^i = loss_{left}^{i-1}/\ \overline{loss^{i-1}}, \quad \lambda_5^i = loss_{bottom}^{i-1}/\ \overline{loss^{i-1}}, \quad \lambda_6^i = loss_{con}^{i-1}/\ \overline{loss^{i-1}}.
\end{aligned}
$$

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.

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

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

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

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

where

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

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

$$(D\cdot\nabla^*)\cdot K\cdot(D\cdot\nabla^*)h = 0, \tag{3.29}$$

where

$$
\begin{aligned}
&(D\cdot\nabla^*)\cdot K\cdot(D\cdot\nabla^*)h = K_{11}[\tfrac{\partial^2}{\partial x^2} - \tfrac{\partial}{\partial x}((\tfrac{z}{\phi-\phi_2}(\phi-\phi_2)_x + \tfrac{\phi_{2,x}}{\phi-\phi_2})\tfrac{\partial}{\partial z}) - (\tfrac{z}{\phi-\phi_2}(\phi-\phi_2)_x \qquad (3.30)\\
&+\tfrac{\phi_{2,x}}{\phi-\phi_2})\tfrac{\partial^2}{\partial z\partial x} + (\tfrac{z}{\phi-\phi_2}(\phi-\phi_2)_x + \tfrac{\phi_{2,x}}{\phi-\phi_2})\tfrac{\partial}{\partial z}((\tfrac{z}{\phi-\phi_2}(\phi-\phi_2)_x + \tfrac{\phi_{2,x}}{\phi-\phi_2})\tfrac{\partial}{\partial z})] + K_{22}\tfrac{1}{(\phi-\phi_2)^2}\tfrac{\partial^2}{\partial z^2} = 0.
\end{aligned}
$$

The boundary conditions of h are given by

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

where $\mathbf{n}=(\pm1,0)$ 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,

$$\frac{\partial h}{\partial x}|_{left,right}=0,\quad \mathbf{n}\cdot K\cdot(D\cdot\nabla^*)h|_{bottom}=0.\tag{3.32}$$

The consistent condition for the boundary conditions is

$$\int_{z=1}\mathbf{n}^{(t)}\cdot K\cdot(D\cdot\nabla^*)h dx=\int_{z=1}[\frac{K_{22}}{\phi-\phi_2}h_z-K_{11}\phi_x(h_x-\frac{\phi_x}{\phi-\phi_2}h_z)]dx=0.\tag{3.33}$$

We define

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

Analogous to the treatment of the top boundary, we expand ϕ_2 in its truncated Fourier Sine series

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

We denote

$$\mathbf{h}_t=(\phi(a),\phi(b),b_1,\cdots,b_n,\phi_2(a),\phi_2(b),c_1,\cdots,c_m).\tag{3.36}$$

The DeepOnet is defined by the following:

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

The loss function is given by

$$\begin{aligned}&L(\theta,\mathbf{h}_t,\epsilon)=\tfrac{1}{n_i}\textstyle\sum_{j=1}^{n_i}[(D\cdot\nabla^*)\cdot K\cdot(D\cdot\nabla^*)G(\mathbf{h}_t,\epsilon,\mathbf{x}_j^t)]^2+\tfrac{1}{n_t}\sum_{j=1}^{n_t}[\phi(\mathbf{x}_j^t)-G(\mathbf{h}_t,\epsilon,\mathbf{x}_j^t)]^2\\&+\tfrac{1}{n_l}\textstyle\sum_{i=1}^{n_l}[G_x(\mathbf{h}_t,\epsilon,\mathbf{x}_i^l)]^2+\tfrac{1}{n_r}\sum_{i=1}^{n_r}[G_x(\mathbf{h}_t,\epsilon,\mathbf{x}_i^r)]^2+\tfrac{1}{n_b}\sum_{i=1}^{n_b}[\tfrac{1}{\phi}G_z(\mathbf{h}_t,\epsilon,\mathbf{x}_i^b)]^2\\&+[\tfrac{\Delta x}{3}(F(\mathbf{x}_1^t,\epsilon)+4\textstyle\sum_{i\ even\in\{2,\cdots,n_t-1\}}F(\mathbf{x}_i^t,\epsilon)+2\sum_{i\ odd\in\{3,\cdots,n_t-2\}}F(\mathbf{x}_i^t,\epsilon)+F(\mathbf{x}_{n_t},\epsilon))]^2.\end{aligned}\tag{3.38}$$

where $\mathbf{x}_j$ are the interior points and $\mathbf{x}_j^k$ 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 used 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 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 the different surface representations produce the same results.

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 given by (3.7) with coefficients or parameters $\mathbf{h}_t$. The sampling of the coefficients in the approximate function 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 $h_0(x) = \sum_{j=1}^{8} b_j \sin(j\pi x)$, we sample $b_1, \cdots b_8$ 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 $b_j := \lambda b_j/h_d, j = 1, \cdots, 8$.
- Then, the top boundary surface is well-represented by vector $\mathbf{h}_t = (\phi(a), \phi(b), b_1, \cdots, b_8)$.

This sampling method makes sure the top boundary fluctuates in $0.5 < \phi(x) \leq 1$. 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 consistent condition that depends on $\mathbf{h}_t$. We denote the input to the neural network in the loss function as follows:

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

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 I number of representing vectors of top boundaries in $\mathbf{h}_t^i, i = 1, \cdots, I$. For each $1 \leq i \leq I$, we have well-defined top boundary $y = \phi_i(x)$. For the ith top boundary, we randomly choose L_i data points $\mathbf{x}_j^{l,i}, j = 1, \cdots, L_i$ on the left boundary, R_i points $\mathbf{x}_j^{r,i}, j = 1, \cdots, R_i$ on the right boundary, B_i points $\mathbf{x}_j^{b,i}, j = 1, \cdots, B_i$ on the bottom, and J_i points in the interior $\mathbf{x}_j^i \in [0, 1] \times [0, \phi(\mathbf{x})], j = 1, \cdots, J_i$, M uniform points $x_j, j = 1, \cdots, M$ in $[0, 1]$.

We divide the dataset into the training and test sets by randomly dividing $\{1, \cdots, I\}$ 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, 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 consistent condition in the loss function.

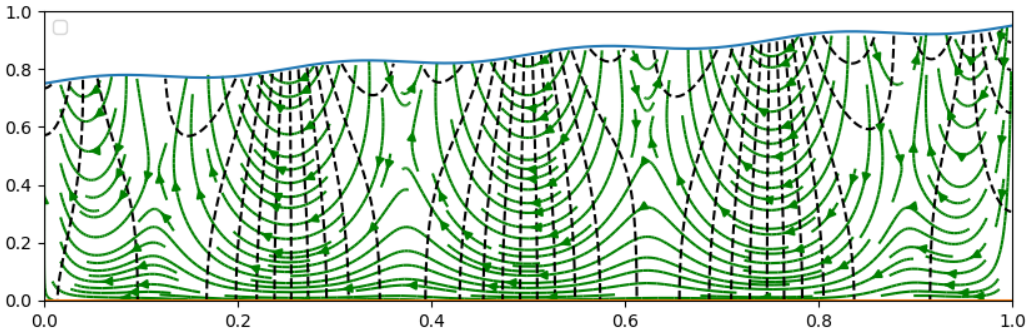
In the DeepONet, the width of branch and trunk net is 200, the depth of the branch net is 4, and the depth of the trunk net is 3. We use the Adam algorithm for the first 1000 epoch optimization step with learning rate 10^{-4} and weight decay 10^{-7}. For the remaining epoch, we use the LBFGS algorithm with learning rate 0.1. 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. All model parameters are summarized in Table 4.1.

Table 4.1: Model Parameters

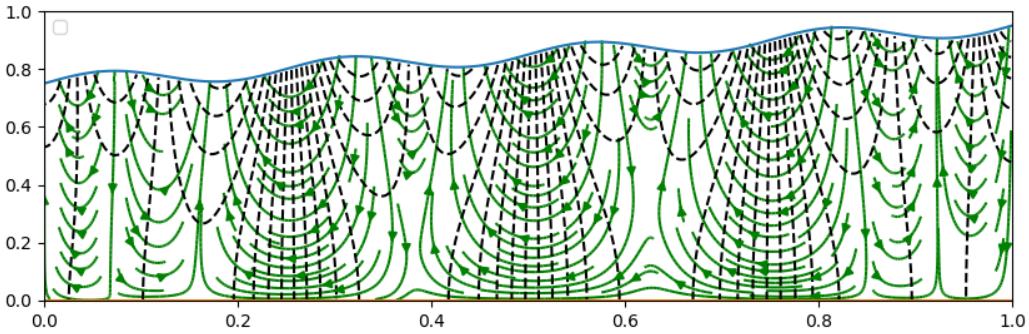
Parameter	Values
Width of trunk net	200
Width of branch net	4
Weight decay	10^{-7}
Learning rate for the first 1000 epoch	10^{-4}
Learning rate for the remaining epoch	0.1
L_x	1000 m
L_y	10000 m
k_{11}	0.01
k_{22}	1

4.1.3 Results obtained using the spectral representation

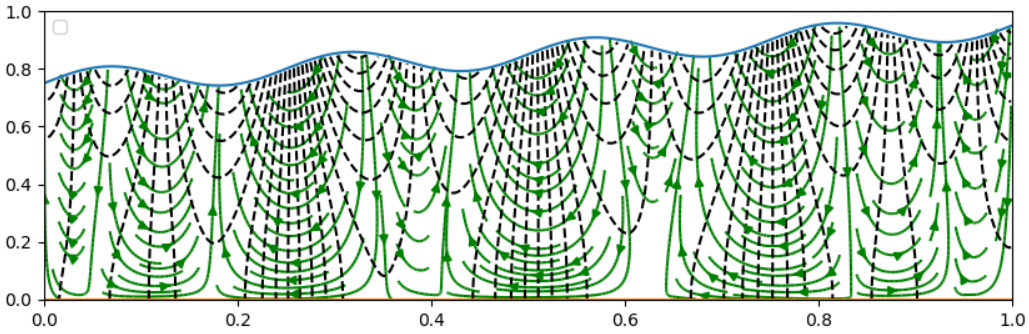
After learning the mapping consisting of the DeepOnet, we present several representative results obtained using the DeepONet to show the steady state solution in the Toth basin. Figure 4.1 depicts the flow field in a Toth basin with sloped top surface topographies of slight variations. There are two factors that impact the flow patterns in the Toth basin: one is the average slope and the other is the local variation of the surface. A larger 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 local variations in the top surface are larger, the long distance transport near the bottom tends to be blocked by localized circulations penetrated downward from the top. Figure 4.1 shows a typical steady state flow pattern due to the surface topography with a fixed average slope and varying local surface variations in space. The two features identified are shown clearly.



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



(b) $\mathbf{h}_t = (0.75, 0.95, 0, 0, 0, 0, 0, 0, 0, 0.03)$.

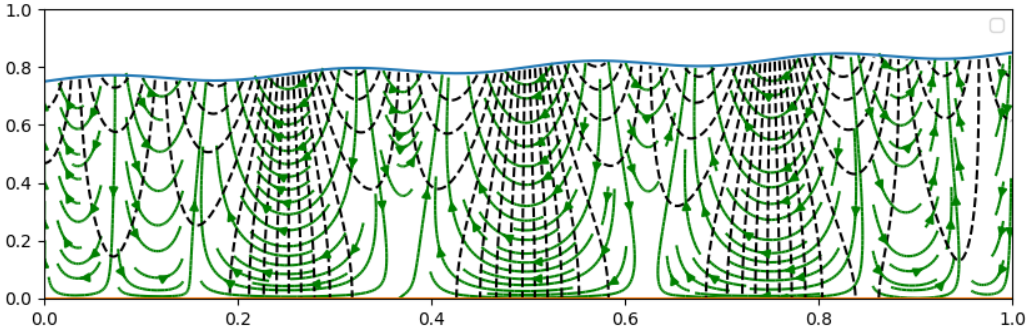


(c) $\mathbf{h}_t = (0.75, 0.95, 0, 0, 0, 0, 0, 0, 0, 0.045)$.

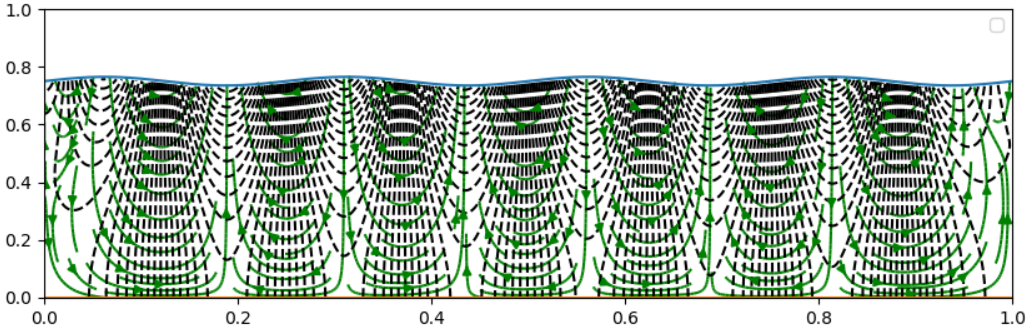
Figure 4.1: The steady state flow patterns in streamlines in the Toth basin with a sloped and varying spatial variations at the top boundary represented by a spectral representation. (a). The slow varying surface topography promotes long distance transport at the bottom of the basin due to the slope. (b). The number of compartmentalized circulations increases when the localized spatial variation in the top is enhanced. But there are still some cross-compartment flows, traveling across two compartments. (c). When spatial variations in space are enhanced further at the top surface, flows are all blocked to compartmentalized circulations so that the long distance transport is halted.

In agreement with Toth's results [2], our solutions also show that the smaller average slope in the topography and small fluctuation in spatial variations apparently promotes compartmentalized

circulations. Figure 4.2 shows two cases of the flow field with smaller average slopes of the top surface. A top surface with an average zero slope but some spatial variation creates several fully compartmentalized flow patterns correlated with the wave form of the top boundary.

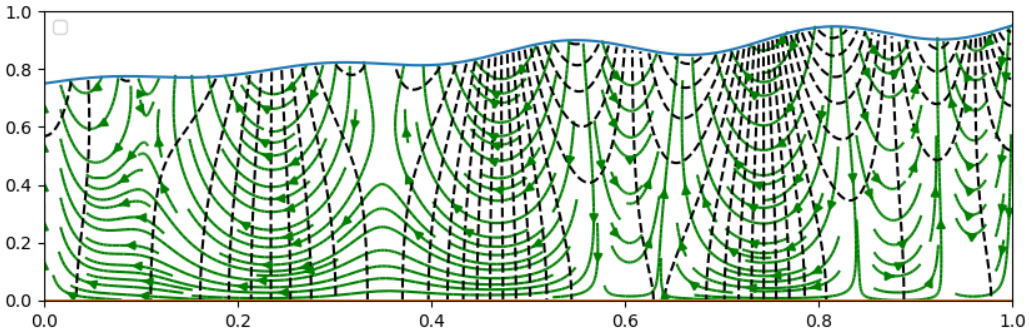


(a) $\mathbf{h}_t = (0.75, 0.85, 0, 0, 0, 0, 0, 0, 0, 0.015)$.

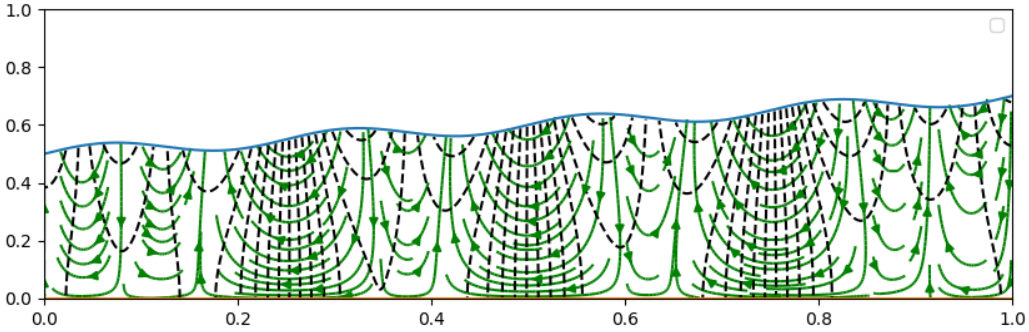


(b) $\mathbf{h}_t = (0.75, 0.75, 0, 0, 0, 0, 0, 0, 0, 0.015)$.

Figure 4.2: Steady state flows with top surface of smaller slopes. Smaller 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)



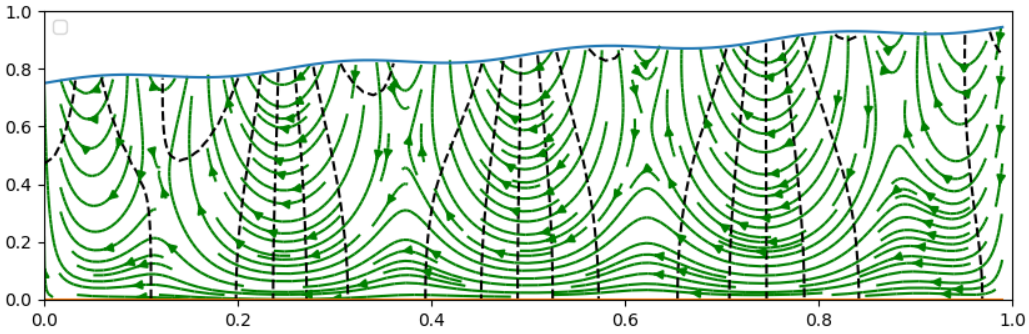
(b)

Figure 4.3: (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 enahnces on the left in this plot. Parameter values are $\mathbf{h}_t = [0.5, 0.7, -1.276 \times 10^{-6}, 2.679 \times 10^{-6}, -4.382 \times 10^{-6}, 6.69 \times 10^{-6}, -1.028 \times 10^{-5}, 1.717 \times 10^{-5}, -3.732 \times 10^{-5}, 0.025]$.

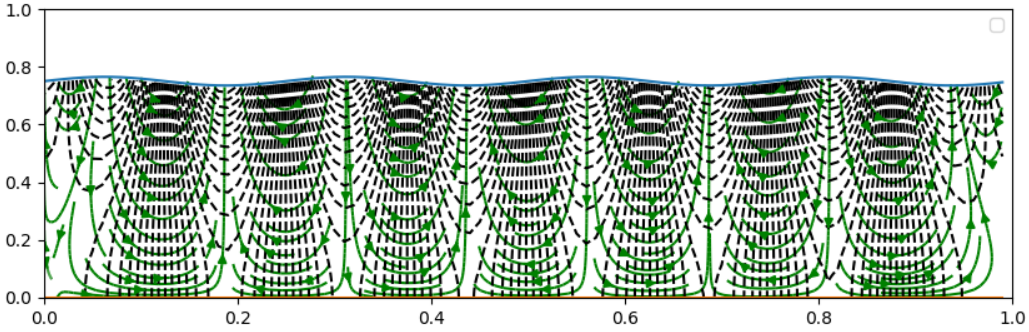
We observe that local curvatures in the top surface affect flow patterns in the Toth basin. The larger local curvature tends to create more localized flow patterns while the smaller one promotes more global flow transport patterns in the bottom of the basin. Figure 4.3 depicts a case where the magnitude of the curvature at the left is smaller than that on the right of the basin. As the result, the flow patten is more localized on the right than on the left of the basin. We note that it is the overall slope of the surface topography that dominates the overall flow pattern, while the local curvature limits the flow pattern to be more localized. 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 Figure 4.3-b since it was an asymptotic study over a near flat top boundary. This study indeed extends the asymptotic analyses in [2] to the truly nonlinear topography of potentially large spatial variations. With the boundary-to-solution mapping given by the DeepONet, we can literally calculate any solution pattern so long as we know the top surface topography.

4.1.4 Results obtained by the domain mapping method

Here we report the results obtained using the domain mapping method on the same two top boundaries. The results are the same as expected. Figure 4.4 shows two calculated flow fields using the DeepOnet obtained from the domain mapping method.



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



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

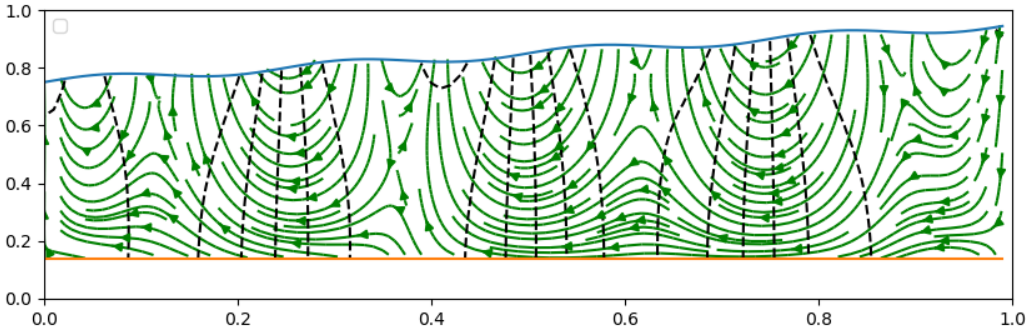
Figure 4.4: Steady state solutions obtained using the domain mapping method. (a). This is identical to the one with the same parameters in Figure 4.1-a. (b). This is identical to the one with the same parameters in Figure 4.2-b. Thus, the two methods produce the same results.

4.2 Results with arbitrary top and bottom boundaries

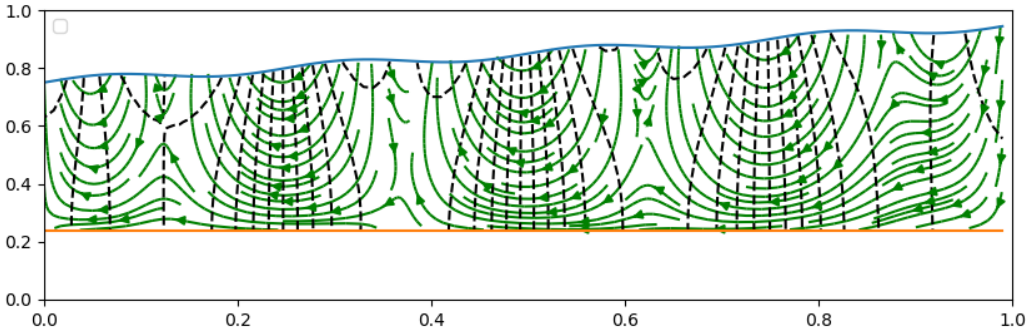
In this paper, we also use the domain mapping method to study the case where the bottom boundary is of an arbitrary shape. The bottom boundary is sampled the same as the top boundary, except that some coefficients/parameters are different. Specifically, $\phi_2(a)$ is sampled uniformly from $[0.14, 0.24]$, $\phi_2(b)$ uniformly from $[\phi(a) - 0.14, \phi(a) + 0.14]$, $\lambda \in [0, 0.1]$. Compared to the case where the bottom boundary is flat, we are interested in two issues here: 1. how does the depth between the top and the bottom boundaries affects the flow pattern in the basin? 2. how does the morphology of the bottom boundary affect the flow field in the basin in addition to that of the top boundary?

When the bottom is flat, a decrease in the depth of the basin does not seem to impact much to the overall flow pattern except that the localized/compartmentalized circulation is enhanced at the top and the depth of the circulation region becomes larger as shown in Figure 4.5. When the flat bottom boundary is inclined in the same direction as the top boundary does, the local circulatory flow seems to increase near the top boundary as shown in Figure 4.5. When the bottom is inclined

opposite to that of the top boundary, the increased depth in the far right end alleviates the small scale circulatory motion to a slightly long distance transport pattern across a scale much larger than the previously confined circulatory region (see Figure 4.5). When the bottom boundary is wavy, it does not seem to add any new features to the already known flow patterns alluded to earlier. Figure 4.7 depicts two cases where the bottom boundaries are wavy with different amplitudes of spatial variations.

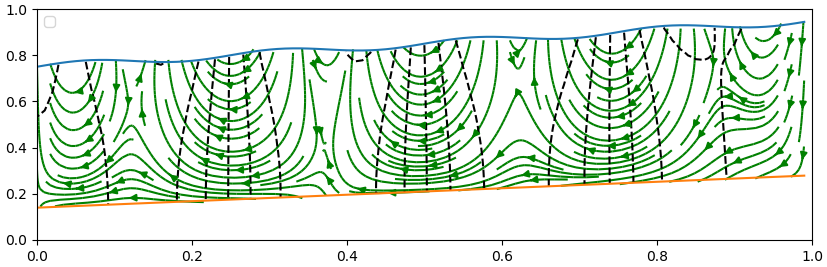


(a) $\mathbf{h}_t = (0.14, 0.14, 0, 0, 0, 0, 0, 0, 0, 0)$.

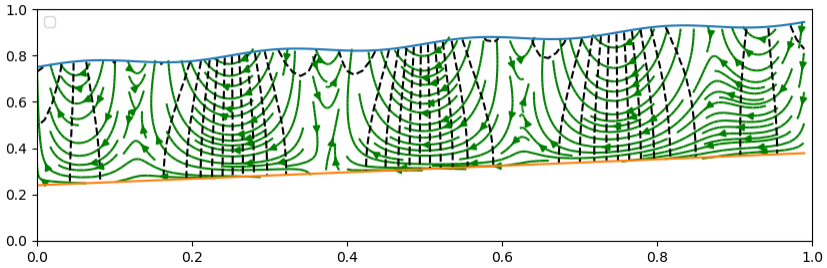


(b) $\mathbf{h}_t = (0.24, 0.24, 0, 0, 0, 0, 0, 0, 0, 0)$.

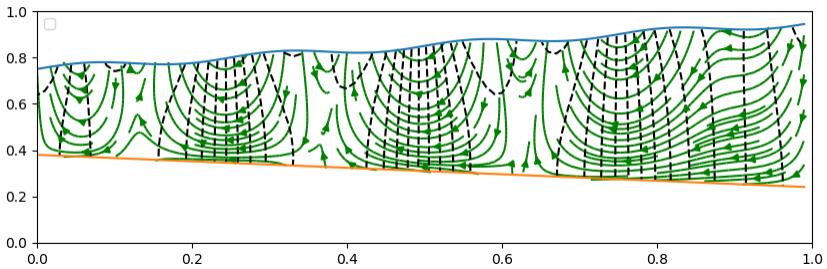
Figure 4.5: The steady state solution in the Toth basin with given top and bottom boundaries represented by spectral representations obtained using the domain mapping approach. (a). The flow pattern is altered compared to Figure 4.1. With the bottom lifted so that the basin becomes shallower, the flows are compressed. (b). With the bottom lifted further, the flows are compressed. However, the flow pattern does not seem to differ from (a) qualitatively.



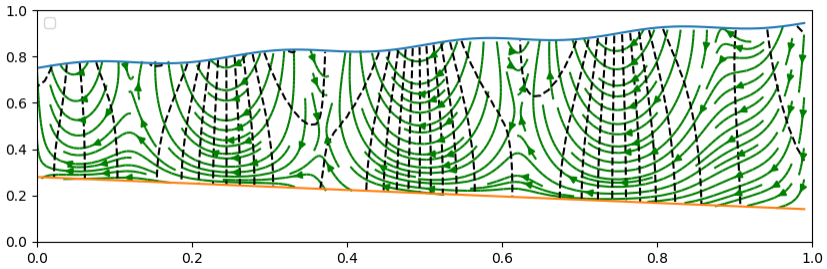
(a) $\mathbf{h}_t = (0.14, 0.28, 0, 0, 0, 0, 0, 0, 0, 0)$.



(b) $\mathbf{h}_t = (0.24, 0.38, 0, 0, 0, 0, 0, 0, 0, 0)$.

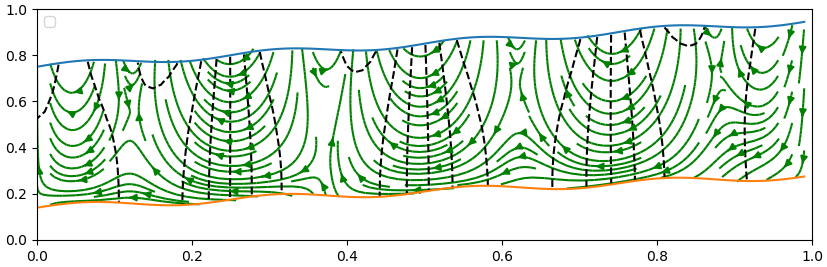


(c) $\mathbf{h}_t = (0.38, 0.24, 0, 0, 0, 0, 0, 0, 0, 0)$.

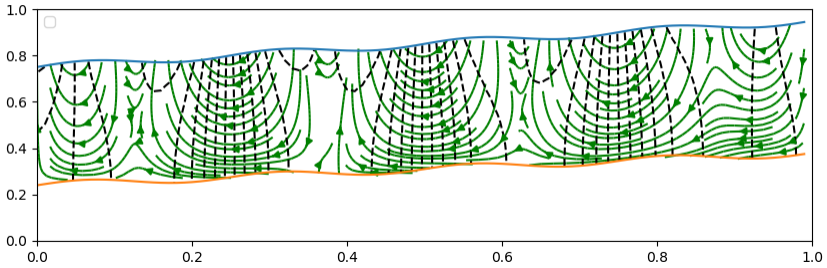


(d) $\mathbf{h}_t = (0.28, 0.14, 0, 0, 0, 0, 0, 0, 0, 0)$.

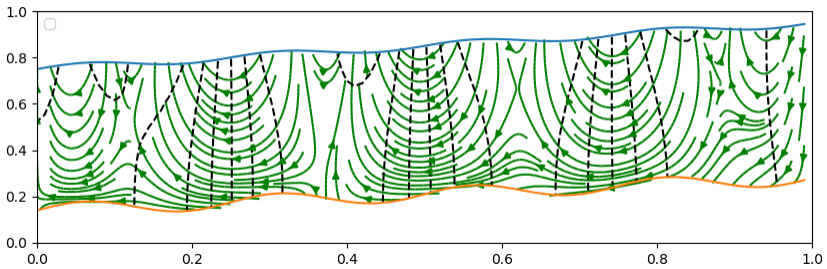
Figure 4.6: The steady state solution in the Toth basin with a given top and bottom boundaries represented by spectral representations obtained using the domain mapping approach. (a). The flow has a tendency to flow to the left in long distance due to the increasing slope over there. (b). The flow tends 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 of transport is observed near the bottom.



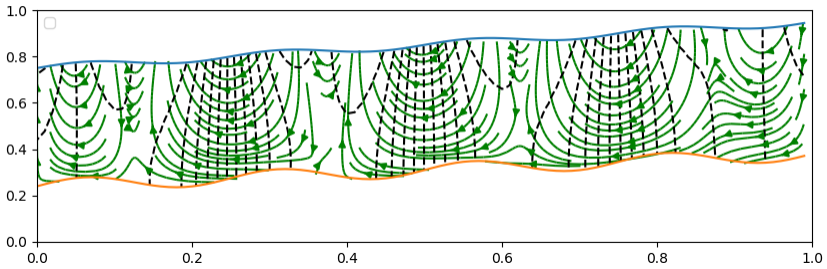
(a) $\mathbf{h}_t = (0.14, 0.28, 0, 0, 0, 0, 0, 0, 0, 0.015)$.



(b) $\mathbf{h}_t = (0.24, 0.38, 0, 0, 0, 0, 0, 0, 0, 0.015)$.



(c) $\mathbf{h}_t = (0.14, 0.28, 0, 0, 0, 0, 0, 0, 0, 0.03)$.



(d) $\mathbf{h}_t = (0.24, 0.38, 0, 0, 0, 0, 0, 0, 0, 0.03)$.

Figure 4.7: The steady solution in the Toth basin with a wavy top and bottom boundaries represented by spectral representations obtained using the domain mapping approach. (a). The flow has a tendency to flow to the trough of the bottom, when bottom boundary isn't flat. (b). A slight decrease in depth does not make much qualitative difference in the flow pattern. (c). The tendency in (a) become more obvious due to the increase in the fluctuating amplitude in the bottom surface. (d) The pattern is similar to that in (b).

4.3 Robin boundary value problem

For the Robin boundary condition given in (2.14), we define the loss function as follows:

$$L(\theta, \mathbf{h}_t, \epsilon) = \frac{1}{n_i}\sum_{j=1}^{n_i}(\nabla \cdot K \cdot \nabla G(\mathbf{h}_t, \epsilon, \mathbf{x}_j))^2 + \frac{1}{n_t}\sum_{j=1}^{n_t}[-\gamma(\phi(\mathbf{x}_j^t) - G(\mathbf{h}_t, \epsilon, \mathbf{x}_j^t))+$$

$$\mathbf{n} \cdot K \cdot \nabla G(\mathbf{h}_t, \epsilon, \mathbf{x}_j^t)]^2 + \frac{1}{n_l}\sum_{i=1}^{n_l}(G_x(\mathbf{h}_t, \epsilon, \mathbf{x}_i^l))^2 + \frac{1}{n_r}\sum_{i=1}^{n_r}(G_x(\mathbf{h}_t, \epsilon, \mathbf{x}_i^r))^2+ \tag{4.2}$$

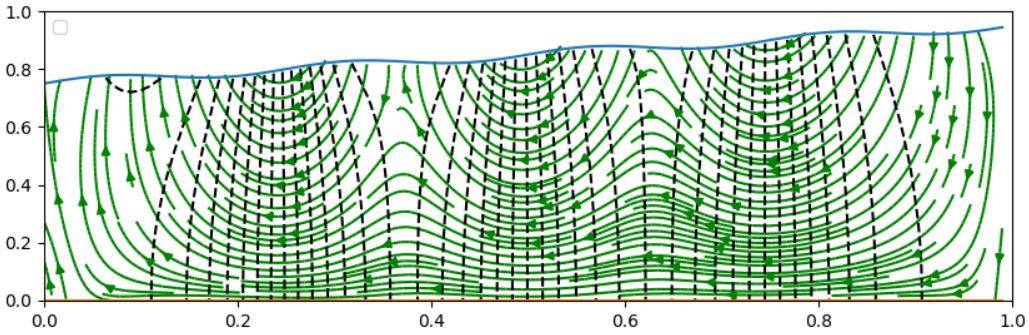
$$\frac{1}{n_b}\sum_{i=1}^{n_b}(G_y(\mathbf{h}_t, \epsilon, \mathbf{x}_i^b))^2 + [\frac{\Delta x}{3}\big(F(\mathbf{x}_1, \epsilon) + 4\sum_{i\ even\in\{2,\cdots,n-1\}} F(\mathbf{x}_i^t, \epsilon)+$$

$$2\sum_{i\ odd\in\{3,\cdots,n-2\}} F(\mathbf{x}_i^t, \epsilon) + F(\mathbf{x}_n^t, \epsilon)\big)]^2,$$

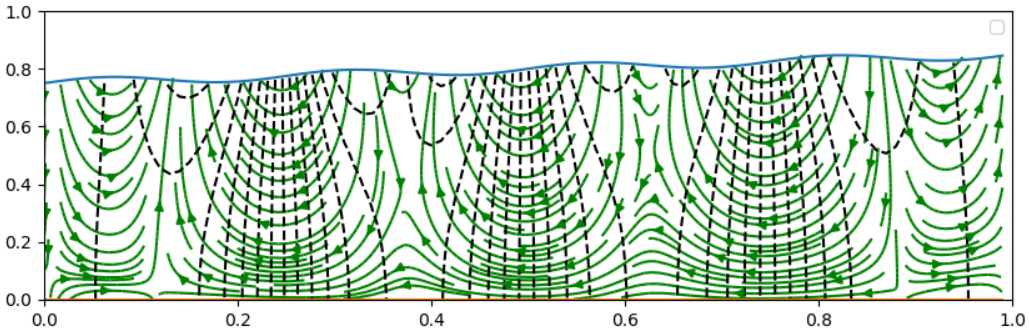
where $\mathbf{x}_j$ are the interior points and $\mathbf{x}_j^k$ are boundary points at the top, left, right, and bottom boundary, respectively, $\mathbf{n} = \frac{1}{\sqrt{1+\phi_x^2}}(-\phi_x, 1)$, and $F(\mathbf{x}, \epsilon) = G(\mathbf{h}_t, \epsilon, \mathbf{x}) - \phi(\mathbf{x})$.

The model parameters are the same as we used above. γ can be identified as the reciprocal of the penetration length: the larger γ is, the smaller its impact to the flow in the bottom of the basin. Hence, it is expected that the flow pattern should be similar to the case of the Dirichlet boundary condition when γ is large. γ is treated as a hyperparameter, which we must set 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 which we decide not to pursue in this study. For the Robin boundary condition, we are interested in what change it brings to the steady state flow patterns in comparison to the previous case.

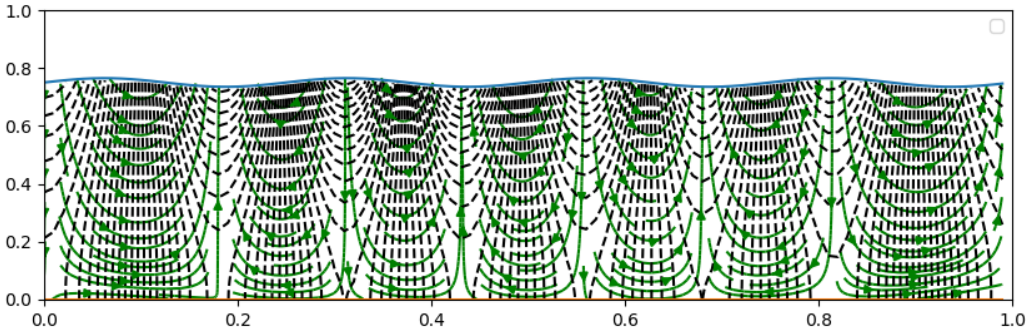
By examining the numerical results, we observe that when $\gamma \geq 10$, the flow patterns in the three cases with different slopes and spatial variations are pretty much independent of the increase in values of γ. Namely, the patterns in Figure 4.8 ($\gamma = 10$) are nearly the same as those in Figure 4.9 when $\gamma = 10^7$. While γ becomes smaller however, the flow patterns change quite dramatically, the number of localized flow patterns descreses and long distance transport becomes more prominent. This is because the penetration length is reduced in this case so that the flow patterns near the bottom are less affected by the top boundary condition. We expect that the solution is going to be a constant as $\gamma \to \infty$. The steady states we have calculated do support the trend.



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

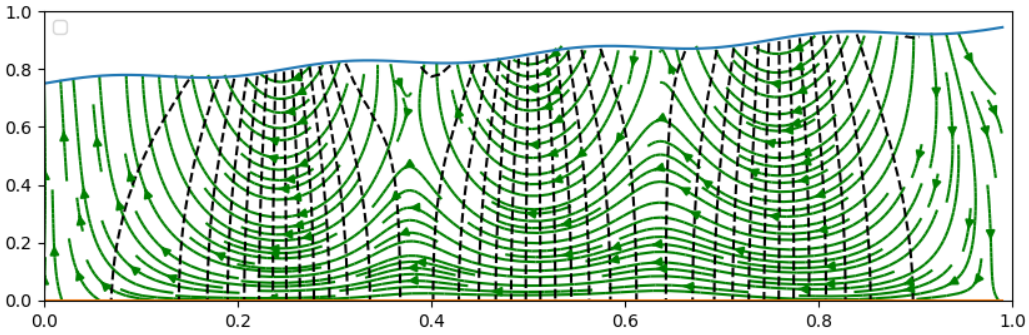


(b) $\mathbf{h}_t = (0.75, 0.85, 0, 0, 0, 0, 0, 0, 0, 0.015)$.

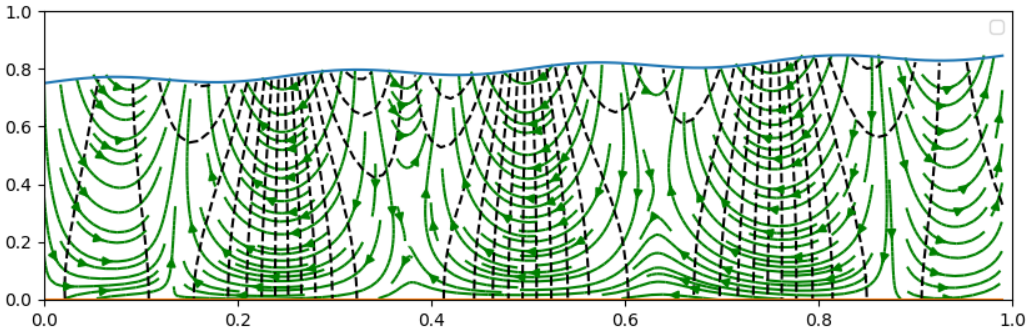


(c) $\mathbf{h}_t = (0.75, 0.75, 0, 0, 0, 0, 0, 0, 0, 0.015)$.

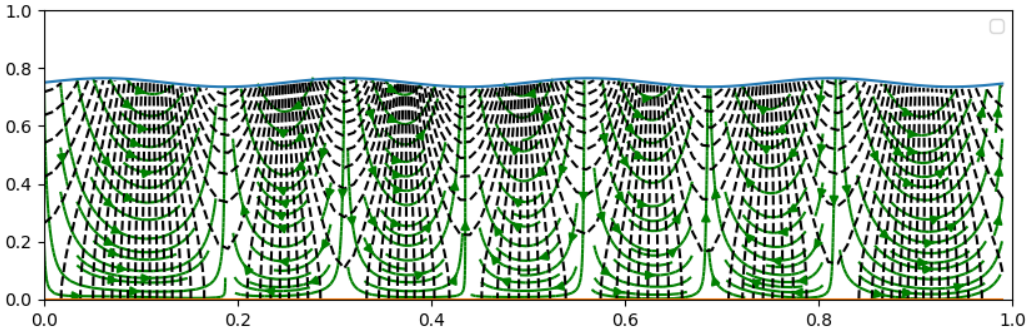
Figure 4.8: Steady state solution patterns at three selected slopes of the top boundary with $\gamma = 10$. (a). The top boundary of a higher slope produces long-distance transport near the impervious bottom boundary. (b). Compartmentalization becomes more prominent as the slope reduces and in the meantime the number of circulatory flow cell increases. (c) When the average slope of the top surface is zero, the long distance transport completely ceases and compartmentalized flow patterns dominates.



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

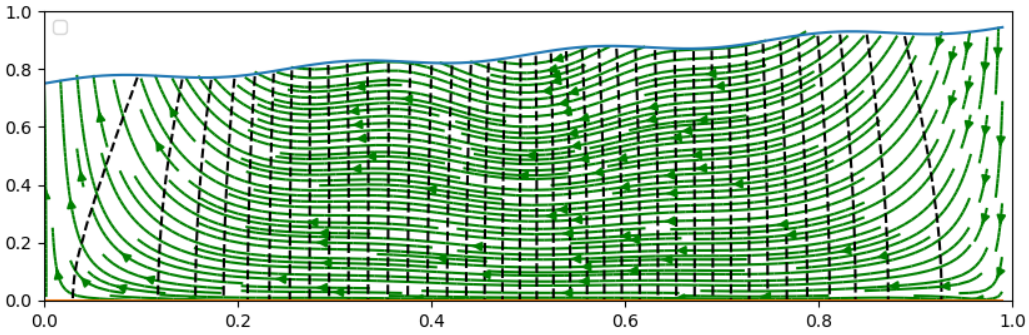


(b) $\mathbf{h}_t = (0.75, 0.85, 0, 0, 0, 0, 0, 0, 0, 0.015)$.

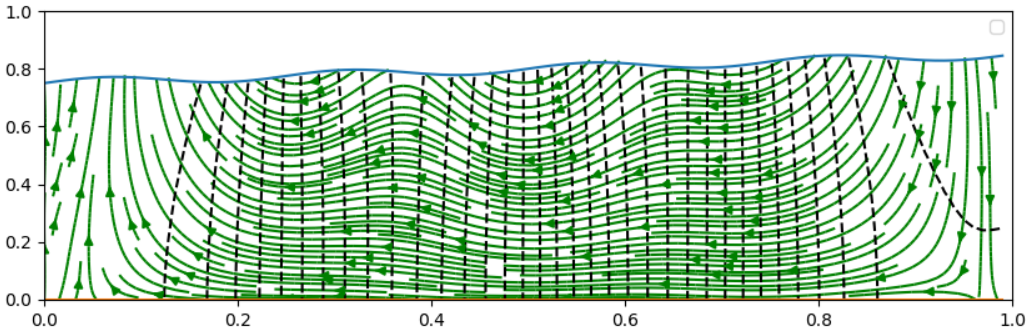


(c) $\mathbf{h}_t = (0.75, 0.75, 0, 0, 0, 0, 0, 0, 0, 0.015)$.

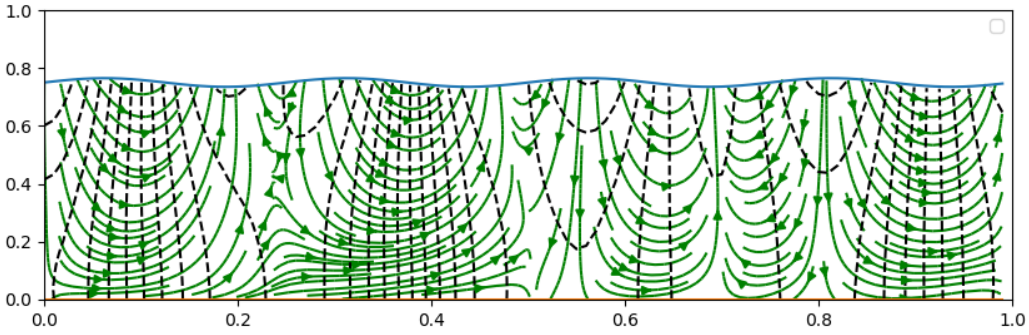
Figure 4.9: Steady state solution patterns at three selected slopes of the top boundary with $\gamma = 10^7$. (a). The top boundary of a higher slope produces long-distance transport near the impervious bottom boundary. (b). Compartmentalization becomes more prominent as the slope reduces and in the meantime the number of circulatory flow cell increases. (c) When the average slope of the top surface is zero, the long distance transport completely ceases and compartmentalized flow patterns dominates. The results are similar to those in Figure 4.8.



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



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



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

Figure 4.10: Steady state solution patterns at three selected slopes of the top boundary with $\gamma = 0.2$. (a). The top boundary of a higher slope produces long-distance transport 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 transport across the entire bottom of the basin ceases. However, the compartmentalized flow cells are larger than in the case where $\gamma \geq 10$.

5 Conclusion

We have developed a novel approach to study the surface topography-steady state flow relation in a Toth basin with arbitrary surface topographies by establishing a boundary to solution mapping

employing the DeepOnet. Once the parameters in the governing transport equation are given, the mapping can be used repeatedly to map out the underground steady state flow field for any given surface topography and the underneath impervious boundary of arbitrary shapes. It provides a novel means to study transport phenomena in complex geophysical systems. In particular, if the geophysical parameters are approximately the same, it can be used to estimate ground water transport patterns in a new Toth basin. Even when applied to other geographical locations where the geophysical parameters change, new models can be machine-learned through transfer learning. With additional computational efforts, the model parameters can also be treated as an input to the DeepOnet so that the mapping can be established for a wider class of transport equations with a range of parameter values. Therefore, the resulting DeepONet can be used as the well-trained "knowledgeable" boundary-to-solution predictor. This perhaps can be viewed as a realization of meta-learning in geophysical applications.

Acknowledgements

Qi Wang's work is partially supported by NSF DMS-1954532 and OIA-1655740, and a GEAR award from SC EPSCoR/IDeA Program. Jun Li's work is partially supported by The Science & Technology Development Fund of Tianjin Education Commission for Higher Education 2020KJ005, and the National Natural Science Foundation of China 11971247.

References

[1] J. Tóth. A theory of groundwater motion in small drainage basins in central Alberta, Canada. *Journal of Geophysical Research*, 67(11):4375–4388, 1962.

[2] J. Tóth. A theoretical analysis of groundwater flow in small drainage basins. *Journal of Geophysical Research*, 68(16):4795–4812, 1963.

[3] Ran An, Xiao Wei Jiang, Jun Zhi Wang, Li Wan, Xu Sheng Wang, and Hailong Li. A theoretical analysis of basin-scale groundwater temperature distribution. *Hydrogeology Journal*, 23(2):397–404, 2014.

[4] M. Bayani Cardenas and Xiao Wei Jiang. Groundwater flow, transport, and residence times through topography-driven basins with exponentially decreasing permeability and porosity. *Water Resources Research*, 46(11):3–23, 2010.

[5] Xu Sheng Wang, L. Wan, Xiao Wei Jiang, H. Li, Y. Zhou, J. Wang, and X. Ji. Identifying three-dimensional nested groundwater flow systems in a tothian basin. *Advances in Water Resources*, 108:139–156, 2017.

[6] Hong Niu, Xing Liang, Sheng Nan Ni, and Zhang Wen. Analytical study of unsteady nested groundwater flow systems. *Mathematical Problems in Engineering*, 2015.

[7] George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.

[8] Lu Lu, Pengzhan Jin, and George Em Karniadakis. Deeponet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators. *ArXiv*, 1910.03193, 2019.

[9] Shuhao Cao. Choose a Transformer: Fourier or Galerkin. *arXiv*, 2105.14995:17, 2021.

[10] John Guibas, Morteza Mardani, Zongyi Li, Andrew Tao, Anima Aanandkumar, and Bryan Catanzaro. Adaptive fourier neural operators: Efficient token mixers for transformers. *arXiv*, 2111.13587:15, 2022.

[11] Georgios Kissas, Jacob Seidman, Leonardo Ferreira Guilhoto, Victor M Preciado, George J Pappas, and Paris Perdikaris. Learning operators with coupled attention. *arXiv*, 2201.01032:39, 2022.

[12] Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier Neural Operator for Parametric Partial Differential Equations. *arXiv*, 2010.08895, 2021.

[13] Guofei Pang, Lu Lu, and George Em Karniadakis. fPINNs: Fractional Physics-Informed Neural Networks. *SIAM Journal on Scientific Computing*, 41(4):A2603–A2626, 2019.

[14] Moshe Leshno, Vladimir Ya. Lin, Allan Pinkus, and Shimon Schocken. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. *Neural Networks*, 6(6):861–867, 1993.

[15] Tianping Chen and Hong Chen. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems. *IEEE Transactions on Neural Networks*, 6(4):911–917, 1995.

[16] Richard Haberman. *Applied Partial Differential Equations: With Fourier Series and Boundary Value Problems*. PEARSON, Boston, 5th ed edition, 2013.

[17] Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3:218–229, 2021.

[18] Guofei Pang, Lu Lu, and George Em Karniadakis. fpinns: Fractional physics-informed neural networks. *SIAM Journal on Scientific Computing*, 41:A2603–A2626, 2019.

[19] Donald A. Nield and Adrian Bejan. *Convection in Porous Media*. Springer, New York, 3rd ed edition, 2006.

[20] J. Toth. A conceptual model of the groundwater regime and the hydrogeologic environment. *Journal of Hydrology*, 10(2):164–176, 1970.

[21] Yakun Li, Wenkai Yu, Jia Zhao, and Qi Wang. Second Order Linear Decoupled Energy Dissipation Rate Preserving Schemes for the Cahn-Hilliard-Extended-Darcy Model. *Journal of Computational Physics*, 444(4):110561, 2021.

[22] Lawrence C. Evans. Partial differential equations, second edition. 2010.

[23] M. Raissi, P. Perdikaris, and G.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.

[24] J. C. Butcher. *Numerical Methods for Ordinary Differential Equations*. Wiley, Chichester, West Sussex, United Kingdom, third edition edition, 2016.