Learning Parameters and Constitutive Relationships with Physics Informed Deep Neural Networks

Alexandre M. Tartakovskya , Carlos Ortiz Marreroa, Paris Perdikarisb,Guzel D. Tartakovskya

, David Barajas-Solanoa

Problems

- physical models of many complex natural systems are “partially” known:

  conservation laws do not provide a closed system of equations.

- Non-homogeneous turbulence, non-Newtonian fluid flow, multiphase flow and transport in porous media, and granular materials:

  accurate theoretical closures are not available.

- Even when sufficiently accurate closed-form partial differential equations(PDE) models are available:

  (space-dependent) parameters are typically unknown.

Method

In the most general case, system dynamics can be described as

$$\frac{\partial u}{\partial t} = F(u)$$

F is a function or differential operator

The following equation describes flow in porous media

$$\frac{\partial u(\mathbf{x},t)}{\partial t} = -\nabla \cdot (K(\mathbf{x},u)\nabla u(\mathbf{x},t))$$

$$K(\mathbf{x},u) = K(\mathbf{x})$$

a linear diffusion equation with heterogeneous diffusion coefficient K(x)

$$K(\mathbf{x},u) = K(u)$$

a nonlinear diffusion equation with a state-dependent coefficient K(u)

Method： physics-informed neural network

differential operator

$$\mathcal{L}[u, K(\mathbf{x}, u)] = 0, \quad \mathbf{x} \in \Omega$$

$$u(\mathbf{x}) = g(\mathbf{x}), \quad \mathbf{x} \in \partial_D \Omega,$$

$$\mathbf{n} \cdot K(\mathbf{x}, u) \nabla u(\mathbf{x}) = q(\mathbf{x}), \quad \mathbf{x} \in \partial_N \Omega.$$

Dirichlet and Neumann boundary conditions

Obviously, we need two neural networks to represent
the state : u() and
the unknown constitutive relationship : K()

$$\hat{u}(\mathbf{x}; \theta) = \mathcal{NN}_u(\mathbf{x}; \theta), \quad \hat{K}(\mathbf{x}, u; \gamma) = \mathcal{NN}_K(\mathbf{x}, u; \gamma)$$

θ and γ are the DNN parameters

Loss function

$$\mathcal{L}[u, K(\mathbf{x}, u)] = 0$$

$$\mathbf{n} \cdot K(\mathbf{x}, u)\nabla u(\mathbf{x}) = q(\mathbf{x})$$

Two extra networks represent the eq(1) and eq(3)

$$f(\mathbf{x};\gamma,\theta) = \mathcal{L}\left[\mathcal{NN}_u(\mathbf{x};\theta), \mathcal{NN}_K\left(\mathbf{x}, \mathcal{NN}_u\left(\mathbf{x};\theta\right);\gamma\right)\right] = \mathcal{NN}_f(\mathbf{x};\theta,\gamma),$$

$$f_N(\mathbf{x};\gamma,\theta) = \mathbf{n}\cdot\mathcal{NN}_K\left(\mathbf{x}, \mathcal{NN}_u(\mathbf{x};\theta);\gamma\right)\nabla\mathcal{NN}_u(\mathbf{x};\theta) = \mathcal{NN}_N(\mathbf{x};\theta,\gamma).$$

The corresponding losses are:

$$\frac{1}{N_c}\sum_{i=1}^{N_c} f(\mathbf{x}_i^c;\gamma,\theta)^2$$

$$\frac{1}{N_N}\sum_{i=1}^{N_N}\left[f_N(\mathbf{x}_i^N;\gamma,\theta) - q_i^*\right]^2$$

Eq(2)'s loss: $u(\mathbf{x}) = g(\mathbf{x})$ $\longrightarrow$ $\frac{1}{N_D}\sum_{i=1}^{N_D}\left[\hat{u}(\mathbf{x}_i^D;\theta) - g_i^*\right]^2$

The two networks' loss functions:

$$\hat{u}(\mathbf{x};\theta) = \mathcal{NN}_u(\mathbf{x};\theta) \longrightarrow \frac{1}{N_u}\sum_{i=1}^{N_u}\left[\hat{u}(\mathbf{x}_i^u;\theta) - u_i^*\right]^2$$

$$\hat{K}(\mathbf{x},u;\gamma) = \mathcal{NN}_K(\mathbf{x},u;\gamma) \longrightarrow \frac{1}{N_K}\sum_{i=1}^{N_K}\left[\hat{K}(\mathbf{x}_i^K,\hat{u}(\mathbf{x}_i^K;\theta);\gamma) - K_i^*\right]^2$$

The final loss function:

$$L(\theta,\gamma)=\frac{1}{N_K}\sum_{i=1}^{N_K}\left[\hat{K}(\mathbf{x}_i^K,\hat{u}(\mathbf{x}_i^K;\theta);\gamma)-K_i^*\right]^2+\frac{1}{N_u}\sum_{i=1}^{N_u}\left[\hat{u}(\mathbf{x}_i^u;\theta)-u_i^*\right]^2$$
$$+\frac{1}{N_D}\sum_{i=1}^{N_D}\left[\hat{u}(\mathbf{x}_i^D;\theta)-g_i^*\right]^2+\frac{1}{N_N}\sum_{i=1}^{N_N}\left[f_N(\mathbf{x}_i^N;\gamma,\theta)-q_i^*\right]^2$$
$$+\frac{1}{N_c}\sum_{i=1}^{N_c}f(\mathbf{x}_i^c;\gamma,\theta)^2.$$

$$(\theta,\gamma)=\arg\min_{\theta,\gamma}L(\theta,\gamma).$$

Optimizer: L-BFGS-B (lower cost than SGD)

Application: a linear diffusion equation

$$\nabla \cdot (K(\mathbf{x}) \nabla u(\mathbf{x})) = 0, \quad \mathbf{x} \equiv (x_1, x_2)^T \in (0,1) \times (0,1)$$

$$u(\mathbf{x}) = 1, \quad x_2 = 0 \quad \text{and} \quad u(\mathbf{x}) = 0, \quad x_2 = 1$$

$$\frac{\partial u(\mathbf{x})}{\partial x_1} = 0 \quad x_1 = \{0, 1\}.$$

Dirichlet and Neumann boundary conditions

$$f(\mathbf{x}; \gamma, \theta) = \nabla \cdot [\mathcal{NN}_K(\mathbf{x}; \gamma) \nabla \mathcal{NN}_u(\mathbf{x}; \theta)] = \mathcal{NN}_f(\mathbf{x}; \theta, \gamma)$$

$$f_N(\mathbf{x}; \theta) = \partial \mathcal{NN}_u(\mathbf{x}; \theta) / \partial x_2 = \mathcal{NN}_N(\mathbf{x}; \theta)$$

This equation describes saturated flow in heterogeneous porous media with hydraulic conductivity K(x)

$$L(\theta,\gamma) = \frac{1}{N_K}\sum_{i=1}^{N_K}\left[\hat{K}(\mathbf{x}_i^K;\gamma) - K_i^*\right]^2 + \frac{1}{N_u}\sum_{i=1}^{N_u}\left[\hat{u}(\mathbf{x}_i^u;\theta) - u_i^*\right]^2$$
$$+ \frac{1}{N_D}\sum_{i=1}^{N_D}\left[\hat{u}(\mathbf{x}_i^D;\theta) - g_i^*\right]^2 + \frac{1}{N_N}\sum_{i=1}^{N_N} f_N(\mathbf{x}_i^N;\gamma,\theta)^2$$
$$+ \frac{1}{N_c}\sum_{i=1}^{N_c} f(\mathbf{x}_i^c;\gamma,\theta)^2.$$

Generating reference data:
K: the Gaussian process (zero mean + C(x,x'))
u: finite volume (FV) method

$$C(\mathbf{x},\mathbf{x}') = \sigma^2 \exp(-\|\mathbf{x}-\mathbf{x}'\|^2/2\lambda^2) \qquad \sigma = 1 \text{ and } \lambda = 0.15$$

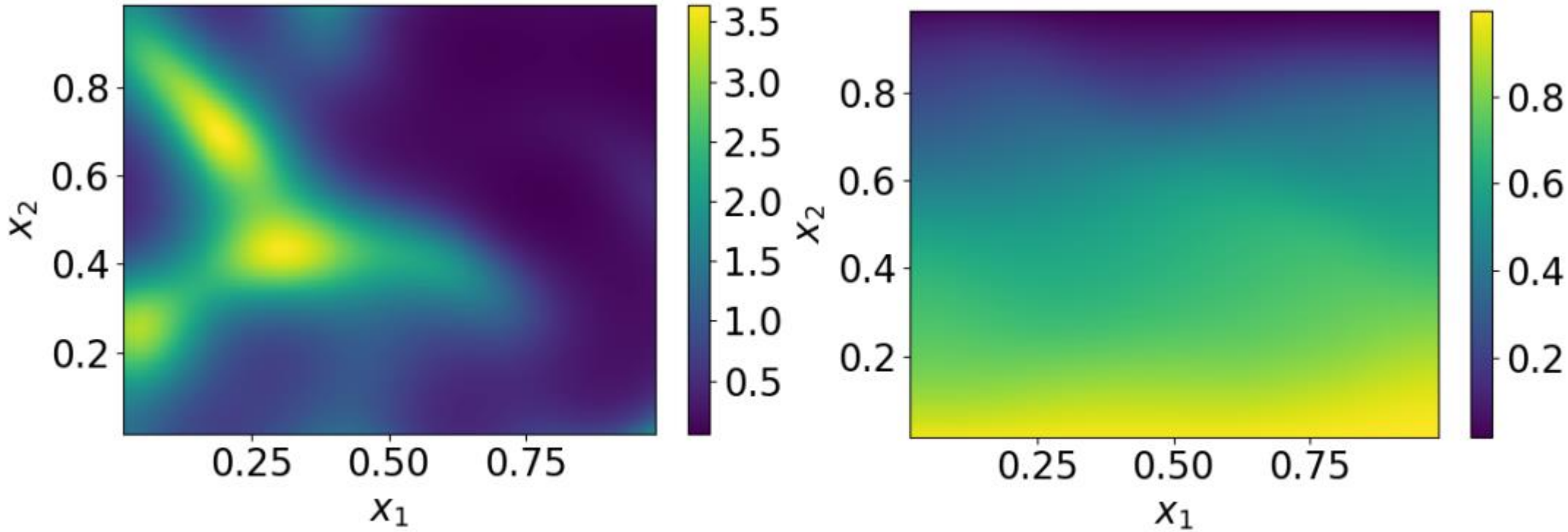
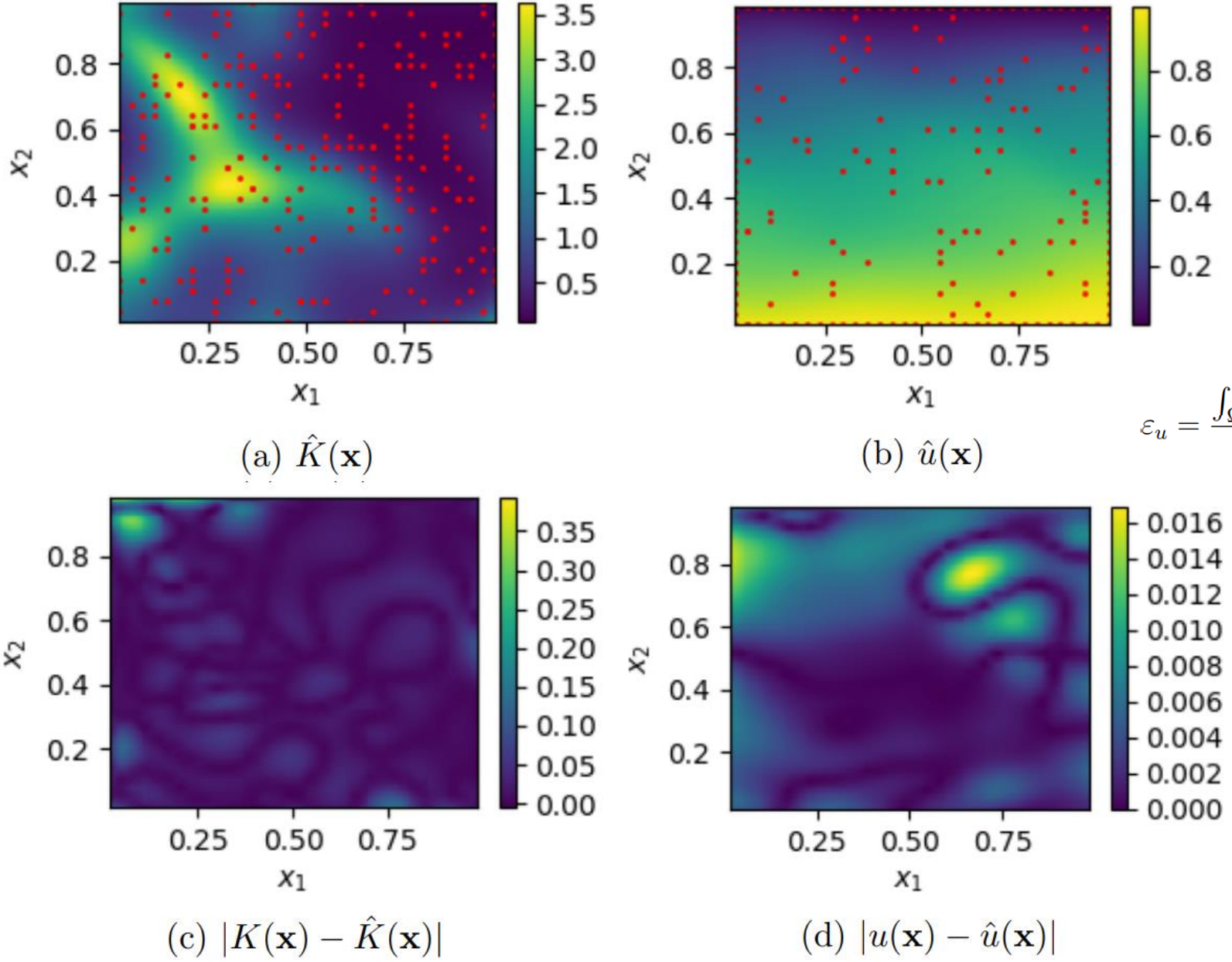


Figure 1: Reference K (left) and u (right) fields.



(a) $\hat{K}(\mathbf{x})$ (b) $\hat{u}(\mathbf{x})$

(c) $|K(\mathbf{x}) - \hat{K}(\mathbf{x})|$ (d) $|u(\mathbf{x}) - \hat{u}(\mathbf{x})|$

$$\varepsilon_u = \frac{\int_\Omega [u(\mathbf{x}) - \hat{u}(\mathbf{x})]^2 \, \mathrm{d}\mathbf{x}}{\int_\Omega u^2(\mathbf{x}) \, \mathrm{d}\mathbf{x}}, \quad \varepsilon_K = \frac{\int_\Omega \left[K(\mathbf{x}) - \hat{K}(\mathbf{x})\right]^2 \, \mathrm{d}\mathbf{x}}{\int_\Omega K^2(\mathbf{x}) \, \mathrm{d}\mathbf{x}}.$$

The relative L2 errors are εu ≈ 0.5% and εK ≈ 1.7 %

Figure 2: Estimated (a) $\hat{K}(\mathbf{x})$ and (b) $\hat{u}(\mathbf{x})$ fields. The red dots indicate the observation locations. Absolute point errors in (c) $\hat{K}(\mathbf{x})$ and (d) $\hat{u}(\mathbf{x})$. $N_K = 250$, $N_u = 100$, and $N_c = 1024$.

study the effect of DNN initialization on the estimated K and u.

$$\overline{\varepsilon}_{(\cdot)} = \frac{1}{N_s}\sum_{i=1}^{N_s}\varepsilon_{(\cdot),i}, \qquad \sigma_{\varepsilon_{(\cdot)}} = \sqrt{\frac{1}{N_s}\sum_{i=1}^{N_s}\left(\varepsilon_{(\cdot),i} - \overline{\varepsilon}_{(\cdot)}\right)^2},$$

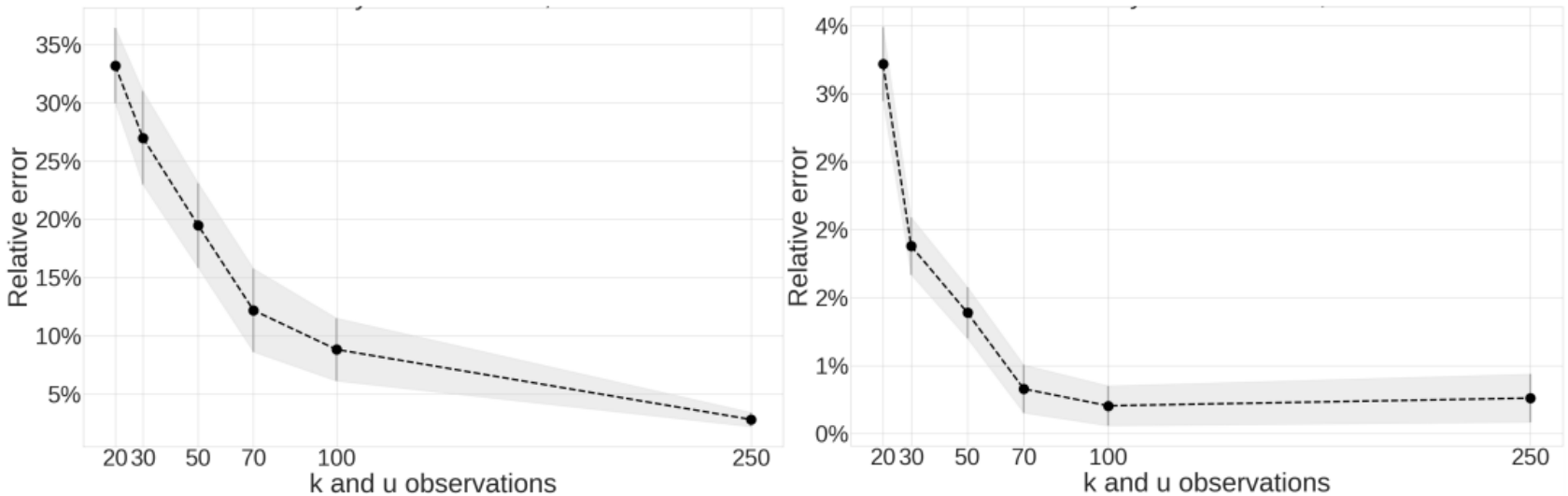
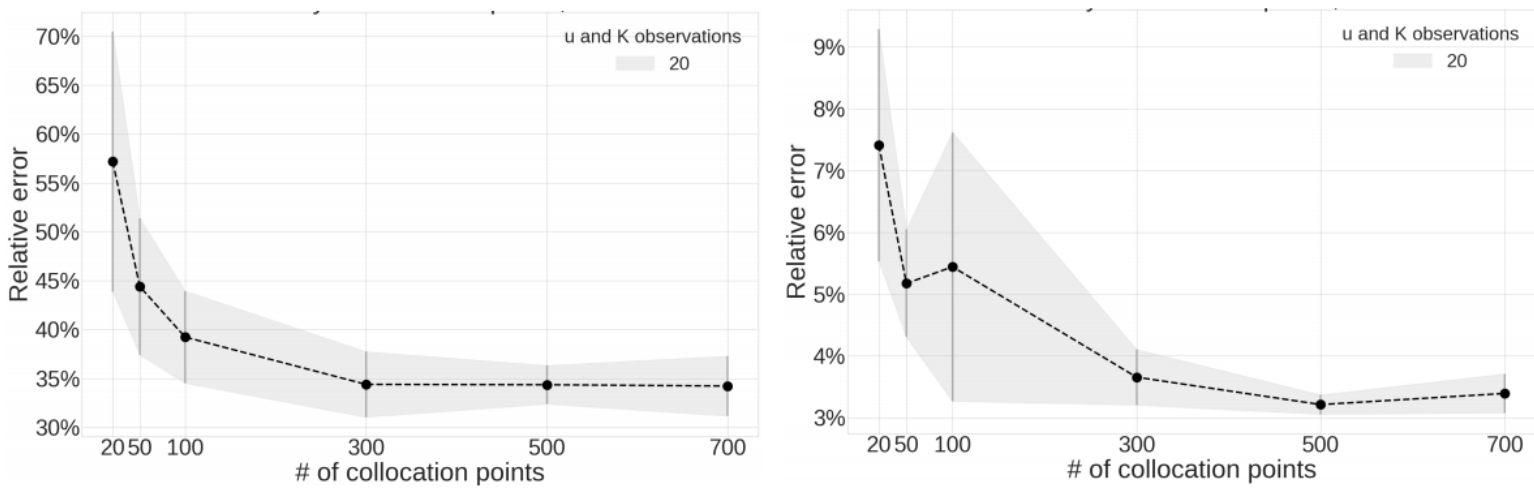


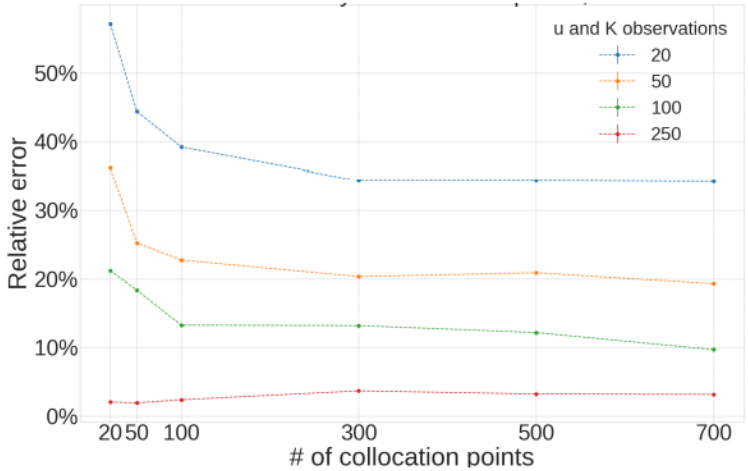
Figure 3: Mean and standard deviation of $\hat{K}$ and $\hat{u}$ obtained with 11 different network initializations using Xavier's initialization scheme as a function of $N = N_K = N_u$. The u and K measurement locations are fixed, and $N_c = 1024$ collocation points are used.

1. uncertainty in ˆu is much smaller than in ˆK
2. the initialization of the DNNs does not have a significant effect on DNN predictions

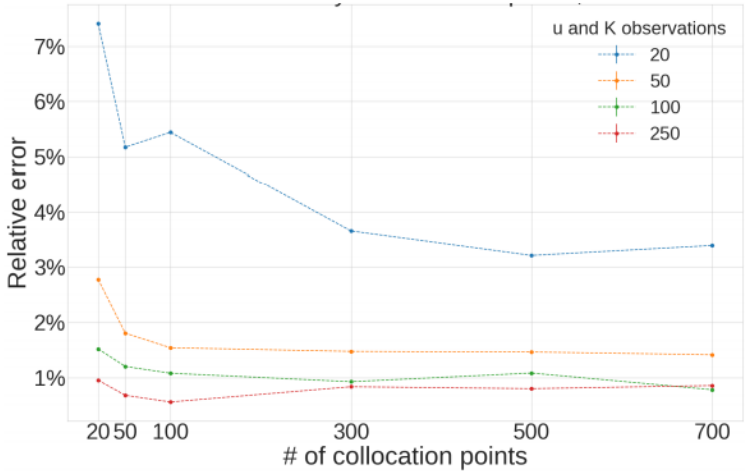


(a) $\overline{\varepsilon}_K$ and σ_{ε_K}

(b) $\overline{\varepsilon}_u$ and σ_{ε_u}



(c) $\overline{\varepsilon}_K$



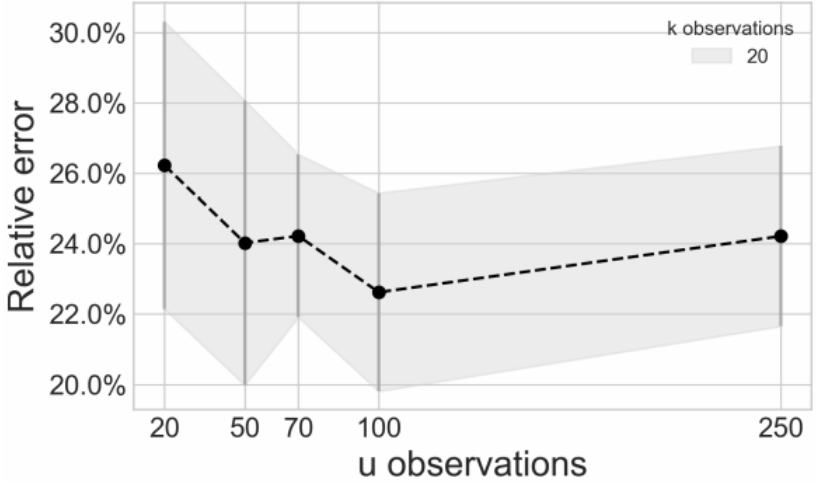
(d) $\overline{\varepsilon}_u$

Figure 4: (a)-(b): Mean and standard deviation of the estimated u and K relative errors as a function of the number of collocation points N_c. For a given N_c, the DNNs are trained 11 times for different configurations of collocation points to compute the mean and variance of the relative errors. The u and K measurement locations are fixed, and $N = N_K = N_u = 20$. Shaded area width is equal to two standard deviations. (c)-(d): Mean K and u relative errors as a function of N_c and N.

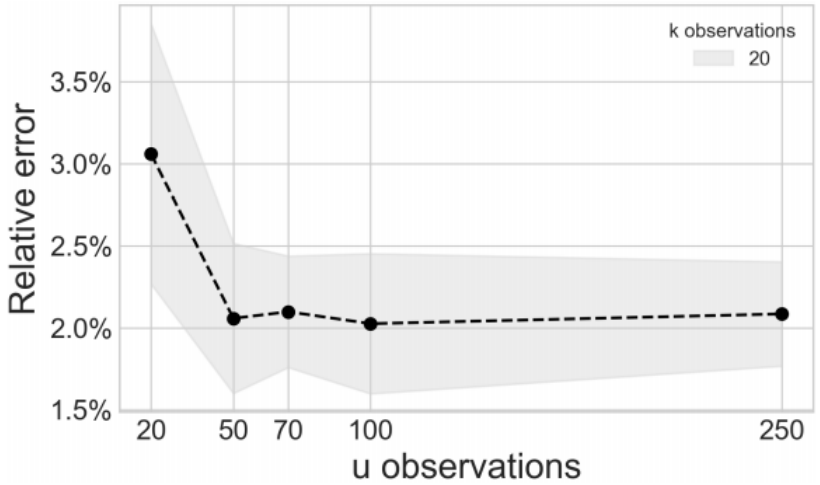
study the effect of the number and location of collocation points

(a),(b): The location of collocation points has a notable effect on the errors, especially for a relatively small Nc

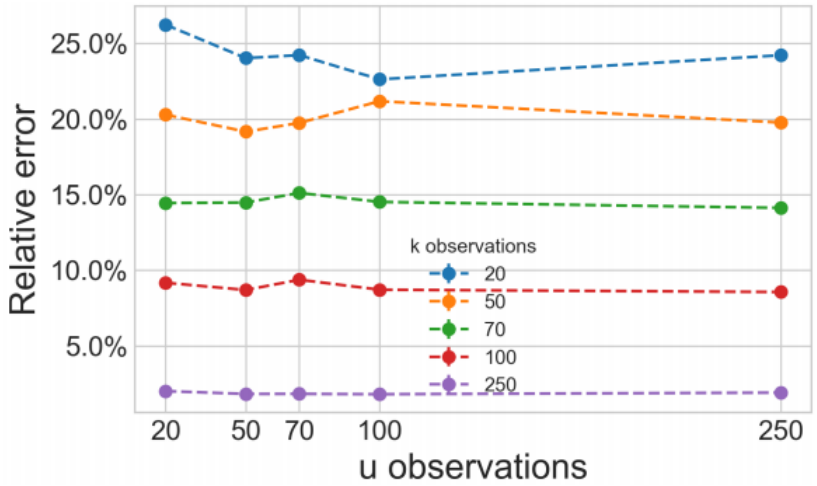
(c),(d): imposing PDE constraints reduces the mean error in K and u by close to 50%



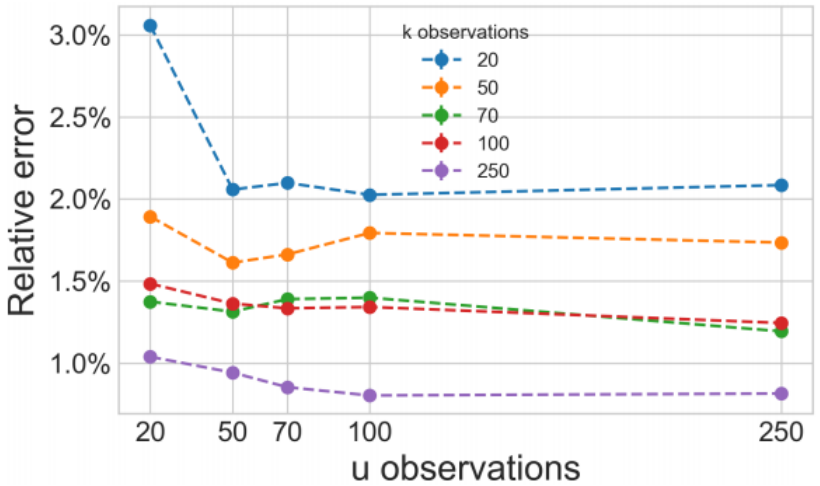
(a) $\overline{\varepsilon}_K$ and σ_{ε_K}



(b) $\overline{\varepsilon}_u$ and σ_{ε_u}



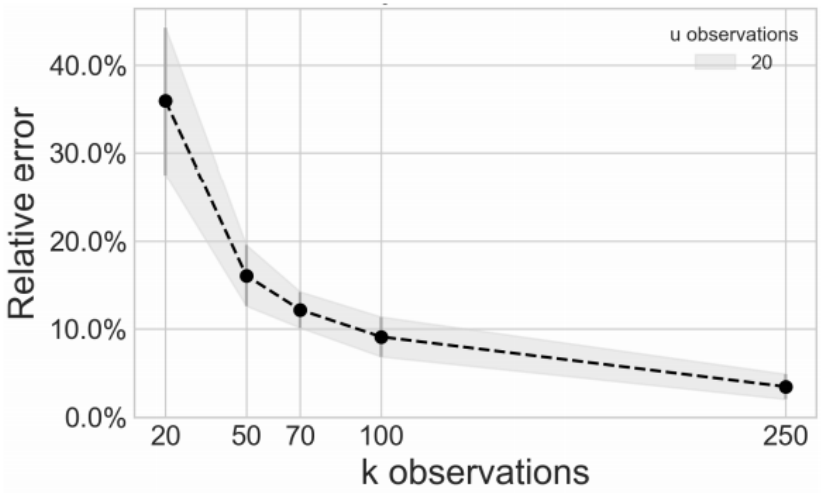
(c) $\overline{\varepsilon}_K$



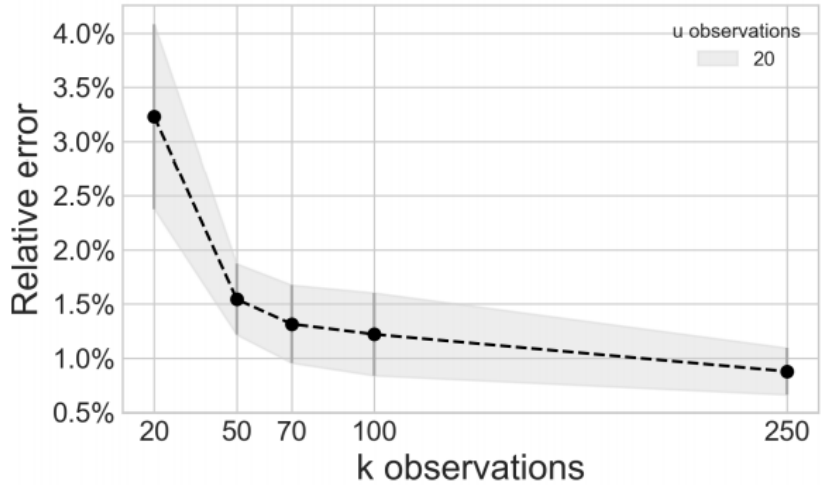
(d) $\overline{\varepsilon}_u$

Figure 5: (a) and (b): Mean relative L_2 error of the predicted K and u as a function of the number of u observations. The number of K observations is 20. Shaded area width is equal to two standard deviations computed for 11 different configurations of u observations. (c) and (d): Mean relative L_2 error of the predicted K and u as a function of the number of u and K observations.

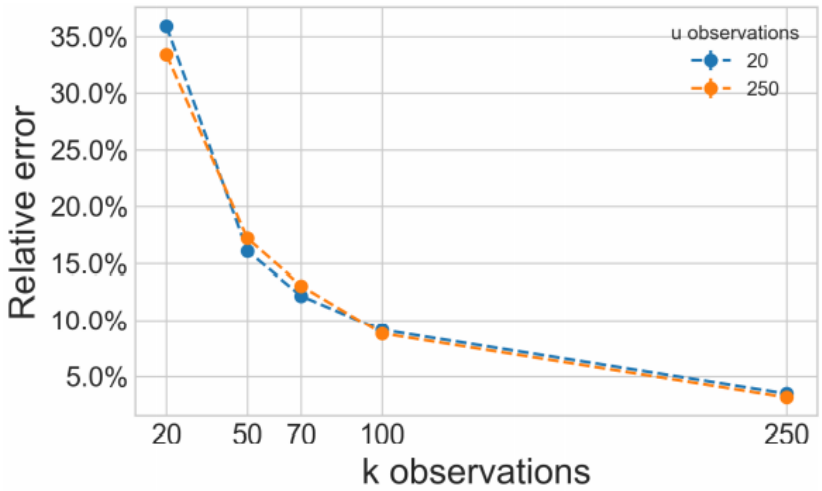
study how the number of K observations versus the number of u observations affects εu and εK



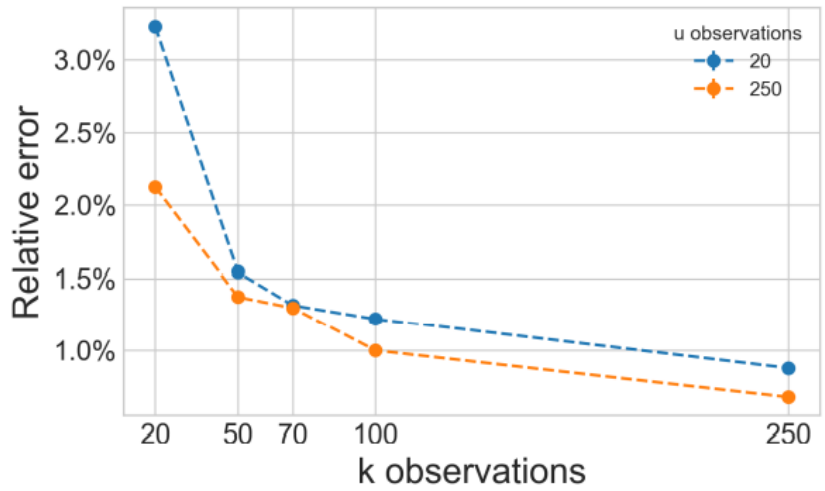
(a) $\overline{\varepsilon}_K$ and σ_{ε_K}



(b) $\overline{\varepsilon}_u$ and σ_{ε_u}



(c) $\overline{\varepsilon}_K$



(d) $\overline{\varepsilon}_u$

Figure 6: (a) and (b): Mean relative L_2 error of the predicted K and u as a function of the number of K observations. The number of u observations is 20. Shaded area width is equal to two standard deviations computed for 11 different configurations of K observations. (c) and (d): Mean relative L_2 error of the predicted K and u as a function of the number of u and K observations.

K measurements are more important than u measurements for reducing error in Kˆ and ˆu.

Compared with MAP

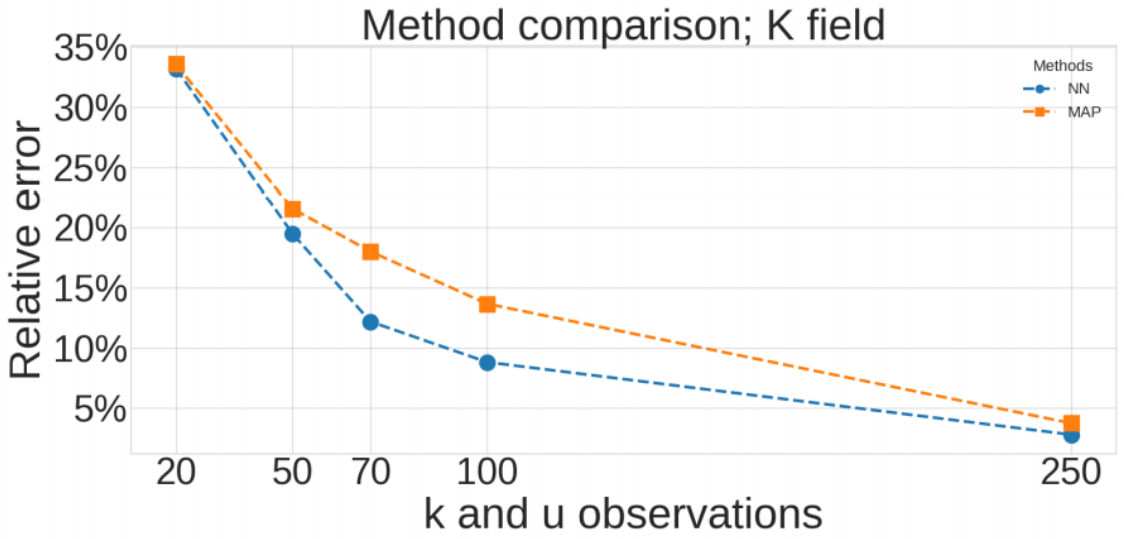


Figure 7: Comparison between L_2 error in the MAP estimate of K and mean L_2 error $\overline{\varepsilon}_K$ in the DNN K estimate.

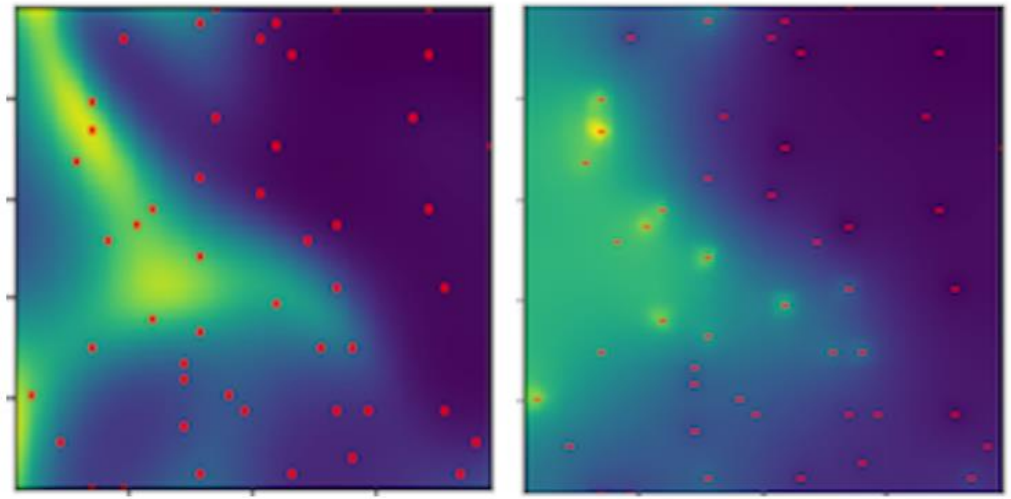


Figure 8: Comparison between the DNN prediction of K (left) and the MAP prediction of K (right) using 50 observations of both u and K.

Application: a nonlinear diffusion equation

$$\nabla \cdot [K(u)\nabla u(\mathbf{x})] = 0, \quad (x_1, x_2) \in (0, L_1) \times (0, L_2)$$

subject to the boundary conditions

$$\begin{aligned} u(\mathbf{x}) &= u_0, & x_1 &= L_1, \\ -K(u)\frac{\partial u(\mathbf{x})}{\partial x_1} &= q, & x_1 &= 0 \\ \frac{\partial u(\mathbf{x})}{\partial x_2} &= 0, & x_2 &= \{0, L_2\}. \end{aligned}$$

1. A two-dimensional horizontal unsaturated flow (flow of water and air) in a homogeneous porous medium
2. u(x) is the water pressure and K(u) is the pressure-dependent partial conductivity of the porous medium
3. no measurements of K(u) are available

$$\hat{u}(\mathbf{x};\theta) = \mathcal{NN}_u(\mathbf{x};\theta), \quad \hat{K}(u;\gamma) = \mathcal{NN}_K(u;\phi)$$

$$f(\mathbf{x};\theta,\gamma) = \nabla \cdot \left[\mathcal{NN}_K\left(\mathcal{NN}_u(\mathbf{x};\theta);\gamma\right) \nabla \mathcal{NN}_u(\mathbf{x};\theta)\right] = \mathcal{NN}_f(\mathbf{x};\theta,\gamma),$$

$$\mathbf{f}_N(\mathbf{x};\phi,\gamma) = -\mathcal{NN}_K\left(\mathcal{NN}_u(\mathbf{x};\theta);\gamma\right) \nabla \mathcal{NN}_u(\mathbf{x};\theta) = \mathcal{NN}_N(\mathbf{x};\theta,\gamma),$$

$$\begin{aligned}
\mathcal{L}(\theta,\gamma) = & \frac{1}{N_u}\sum_{i=1}^{N_u}\left[\hat{u}(\mathbf{x}_i^u;\theta) - u_i^*\right]^2 \\
& + \frac{1}{N_c}\sum_{i=1}^{N_c} f(\mathbf{x}_i;\theta,\gamma)^2 + \frac{1}{N_D}\sum_{i=1}^{N_D}\left[\hat{u}(\mathbf{x}_i^D;\theta) - u_0\right]^2 \\
& + \frac{1}{N_N^{(x_1)}}\sum_{i=1}^{N_N}\left[f_N^{(x_1)}(\mathbf{x}_i^{N,x_1};\phi,\gamma) - q\right]^2 + \frac{1}{N_N^{(x_2)}}\sum_{i=1}^{N_N}\left[f_N^{(x_2)}(\mathbf{x}_i^{N,x_2};\phi,\gamma)\right]^2
\end{aligned}$$

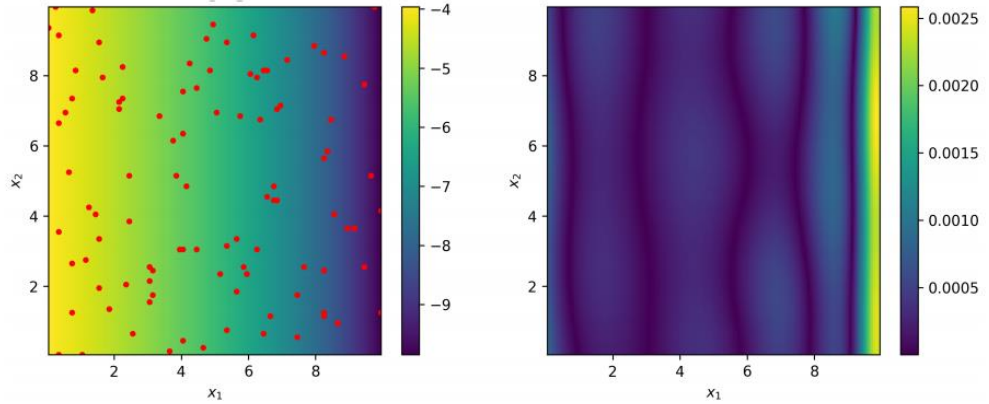


Figure 9: (Left) Referenced $u(\mathbf{x})$ field generated using STOMP with the van Genuchten model for $K(u)$ and the locations of u observations. (Right) Absolute error in the $u(\mathbf{x})$ field estimated with the physics informed DNN.

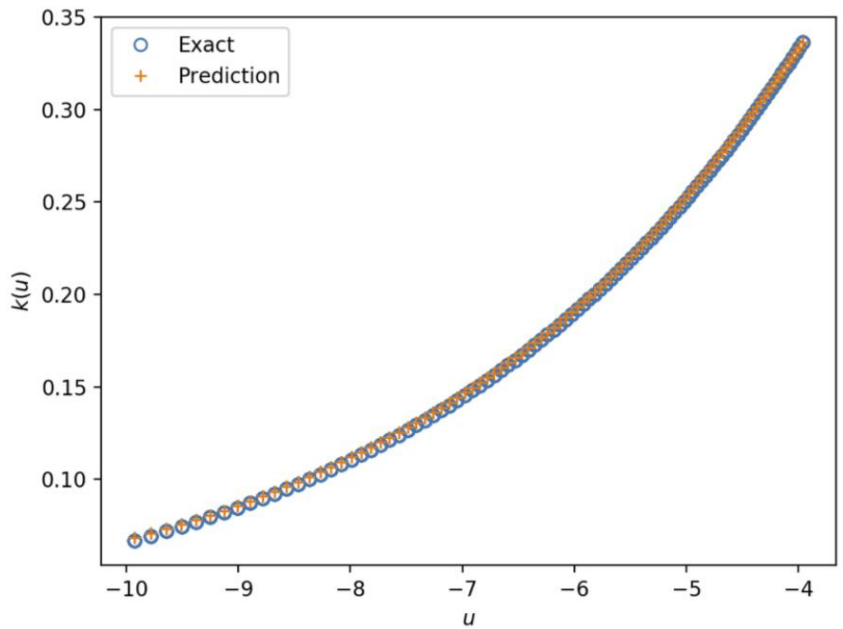


Figure 10: Comparison of the estimated $\hat{K}(u)$ and the reference $K(u)$ given by the van Genuchten model.