Physics-informed machine learning of flow and transport problems

Hongkyu Yoon
Geoscience Research & Applications, Sandia National Laboratories, Albuquerque, NM, USA

Jennifer L. Harding
Geoscience Research & Applications, Sandia National Laboratories, Albuquerque, NM, USA

ABSTRACT: Recent advances in machine learning/deep learning (ML/DL) methods show promising results to enhance our ability to develop fast surrogate models and estimate heterogeneous subsurface properties through inverse modeling approaches. In this work we explore physics-informed neural networks (PINNs) as a way to incorporate governing partial differential equations (PDEs) in the ML framework through loss functions. With advection-dispersion-reaction (ADR) and Darcy equations, PINN methods are evaluated for multiple cases by changing model parameters and the inverse modeling framework. This work demonstrates that PINNs tend to perform better than data-driven only model with less collocation points and the potential capability of PINNs for accurate surrogate models to coupled geomechanical problems in the subsurface.

Keywords: Machine learning, Physics-informed neural network, Inverse modeling, Advection-dispersion-reaction.

1 INTRODUCTION

Recent efforts to apply machine/deep learning (ML/DL) driven models for problems in flow, transport, and coupled thermal-mechanical processes in natural and engineered environments demonstrate promising outcomes (e.g., Kadeethum et al. 2022). Application for subsurface problems is mostly based on a few DL architectures such as convolutional neural networks (CNNs), recursive neural networks, generative adversarial networks (GANs), and etc. “Deep” layers of neural network (NN) with tunable weight and bias parameters can extract features from complex, nonlinear data extremely well, identify the relationships between input parameters/data generation processes (seismic event detection, anomaly detection, segmentation among others), and reconstruct physics of interest (e.g., pressure, stress, velocity). However, these supervised DL methods require large, labelled datasets to expand the generalization of trained model application.

As opposed to purely data-driven DL methods, physics-informed machine learning (PIML) was developed as a promising avenue for advancing engineering and science research by implementing physics-based information into ML process. The PIML framework aims to account for governing equations (e.g., partial differential equations (PDEs)), physical constraints, real measurement data in

learning process (e.g., Raissi et al. 2019; Karniadakis et al. 2021). One of the most popular PIML methods is physics-informed neural networks (PINN) that use NNs to encode governing equations with initial and boundary conditions (IC/BC) as loss functions. After Raissi et al. (2019), this novel method has been applied in almost any science and engineering domains. In addition, the NN training has both a data fit component and reduces a PDE residual, hence PINN can be used as an inverse modeling framework if observed data is available (e.g., Cai et al., 2022).

Compared to traditional numerical methods, for PINNs computations are performed at exact coordinates, hence are independent from the mesh/grid quality. Moreover, the loss term construction that combines the PDE with observed data is trivial, providing a very powerful inverse modeling framework, especially for ill-posed problems (Cai et al., 2021). Compared to standard NN-based models, PINNs can achieve good accuracy with a small number of training data or even without labelled data altogether (Raissi et al. 2019). However, PINNs with fully connected NNs have convergence issues, especially in strong non-linear problems, and have the limit in generalization with IC and BC of PDEs (Wang et al., 2022). Another form of PIML is from a concept of neural operators (NOs)that facilitate a data-driven approach to learn functions from inputs to outputs. As in PINNs, governing equations can be incorporated into the loss functions to develop physics-informed neural operators. There are multiple variants of this class of NOs available such as deep operator network (DeepONet, Lu et al. 2021) and Fourier neural operator (FNO, Li et al. 2020). Although neural operators are very promising, we focus on PINNs in the current work.

Here we use PINNs as forward and inverse models to solve for an advection-dispersion-reaction (ADR) equation and a Darcy equation to demonstrate how PINNs can be used for various conditions in terms of data availability and a set of governing equations. Although we use a coupled flow and transport problem, the methodology can be generalized for many subsurface geomechanics problems.

2 METHODS

In this work we follow a standard PINN framework with multiple governing equations and IC/BC, similar to a multiphysics-informed neural network (MPINN, He et al. 2020). Instead of steady-state advection-dispersion problems in He et al. (2020), we use transient advection-dispersion-reaction (ADR) equation. The Darcy equation and BCs used in this work are identical to the work in He et al. (2020). Due to the PDE of ADR, we set the IC to be zero concentration in the entire model domain. Figure 1 shows an overall schematic of the PINN used in this work. Fully connected feed-forward neural networks (so called dense neural network (DNN)) are used to estimate quantities of interest, which can be concentration, head, and/or hydraulic conductivity in this work.

In general, the workflow of PINN is the following: (1) approximate quantities of interest using DNNs (outputs in Figure 1), (2) compute all derivatives of outputs with respect to space ($\boldsymbol{x}$) and time($\boldsymbol{t}$) using automatic differentiation, (3) compute the loss terms of governing equations, IC, and BCs at a set of points del domain referred as collocation points, and (4) evaluate the mean square error (MSE) and additional regularization terms as the loss term and data reconstruction error if observed or synthetic data are available. This procedure is repeated through optimization of neural networks (i.e., weights and biases) by minimizing the total loss. Since this optimization attempts to minimize the residual form of governing equations, IC, and BCs, the final solution can learn physics of interests. Since the optimization formula is the same as in the inverse modeling, which minimizes the difference between observed data and predictions, we can use the PINN as an inverse model when we have observed data. In this case, the loss term of data ($\boldsymbol{L}_{data}$) is also minimized to estimate the parameter(s) such as hydraulic conductivity given the observed concentrations. The overall PINN architecture and optimization process in this work follows those in He et al. (2020), wherein detailed neural network architectures and definition of terms in the MPINN are provided. Since we work on transient ADR, the addition of IC creates another challenge due to the change of concentrations along inlet boundary at $\boldsymbol{t} > 0$. Figure 2(a) shows the 2-D model domain with BCs. For simplicity, we use an analytical solution provided in Paladino et al. (2018) to generate training/testing datasets.

In this work, we use 4 hidden layers with 40 nodes per each hidden layer using mean square error loss and 10,000 epochs. The model domain has dimensions of 26 x 26 with 11 temporal points. Three cases are evaluated: (1) Case 1 with concentration data-driven only (i.e., no PDEs), (2) Case 2 with

concentration data and ADR, and (3) Case 3 with concentration and head data and ADR and Darcy equations. We also evaluate three different decay values (0, 0.1, and 0.5). In addition, the optimal number of collocation points were determined where governing equations are evaluated. The number of collocation points are selected randomly in space and time, which are training data points, and the rest of the entire field are used as testing set to evaluate the accuracy of trained PINN models.

3 RESULTS AND DISCUSSION

We first evaluate the sensitivity of the PINN relative errors to the number of collocation points

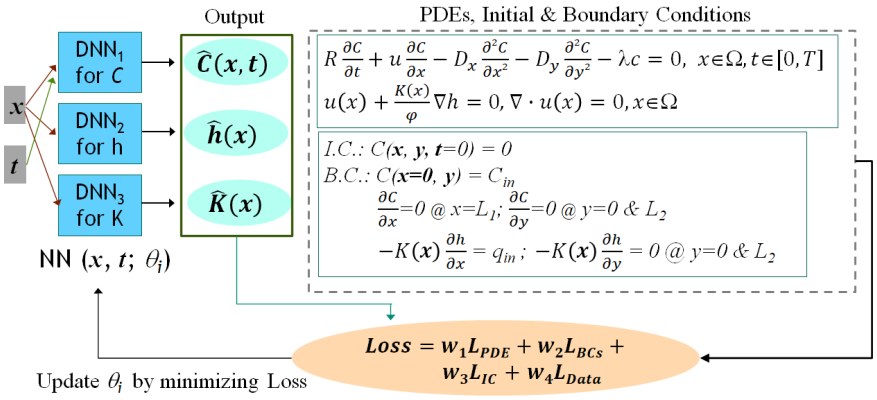

Figure 1. A schematic of a PINN for advection-dispersion-reaction and Darcy equations. I.C. and B.C. stand for initial condition and boundary conditions, respectively. Governing equations of two-dimensional advection-dispersion-reaction and Darcy equations with initial and boundary conditions are shown. A total loss is the sum of four loss terms from residual equations ($\boldsymbol{L}_{PDE}$), initial and boundary conditions ($\boldsymbol{L}_{IC}$, $\boldsymbol{L}_{BCs}$), and data reconstruction ($\boldsymbol{L}_{Data}$) if observed data is available. Weight factors ($\boldsymbol{w}_{1,2,3,4}$) are adjusted to balance the total loss to avoid dominating any one loss term. DNN stands for deep neural network and output represents predicted quantities from DNN. θ_i represent the weights and biases of neural networks where i = concentration (C), head (h), and hydraulic conductivity (K).

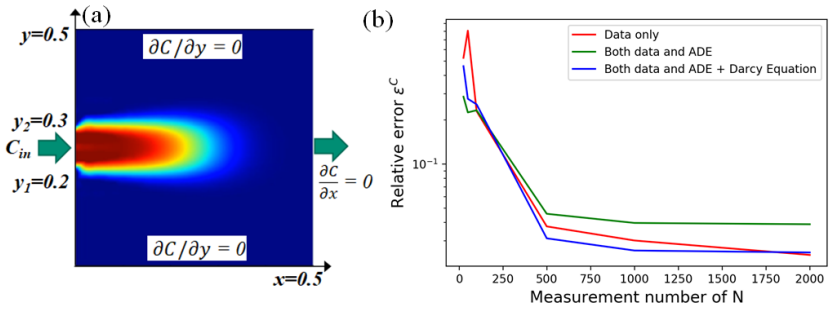

Figure 2. (a) A schematic of two-dimensional model domain (0.5 x 0.5). The inlet concentration ($\boldsymbol{C}_{\text{in}}$) is defined over the central part (0.2<y<0.3) of the inlet boundary with a constant flow velocity ($\boldsymbol{u}$). The blue line shows the advective front of the concentration at time, t. Contour plots are examples of concentration profiles based on analytical solution of ADR in Paladino et al. (2018). (b) Effect of a number of collocation points on the relative error of PINN for three Cases.

(N) for all three Cases. Figure 2(b) shows that the relative error decreases with increasing the number of collocation points (both space and time). Given that the relative error is relatively similar at N =500 and more, we use N = 500 for the rest of cases to save computational time. Accuracy of three different cases with different decay coefficients is evaluated in Table 1. MSE values are computed at collocation points as well as for the entire modeling domain and for all time steps. For N=500, MSE values at collocation points were relatively good at a scale of 10^{-4}~10^{-5}, while the MSE values

at the non-collection grid points increase to the order of 10^{-3}. Interestingly, for N=1,000, MSE values at collocation points are at the order of 10^{-3} and MSE values at the entire domain decrease by ~ one order of magnitude, implying that training at more collocation points can improve the accuracy significantly with additional computational cost. For N=500, Cases 2 and 3 with PINNs perform better than data only Case 1, however, for N=1,000, Case 1 has a better result. It was also reported in He et al. (2020) who showed that N≥80, data-driven only ML performed better than PINNs.

This is mainly due to the fact that in this example a large number of data (N=1,000) are sufficient to train the data-driven model. However, for cases with data availability much smaller as in more complex and realistic conditions, PINNs will perform better than data-driven only models.

Figure 3 shows concentration profiles (ground truth and PINN result) for Case 3 and corresponding loss plots from different terms. Error is overall higher around plume boundary due to fewer collocation points, and, as seen in Table 1, the increase of collocation points improves prediction accuracy. More interestingly, the loss plots in Figure 3(b) show that the overall loss is dominantly governed by the losses from concentration data, the IC, and BCs. This demonstrates learning IC and BCs is not an easy task given the importance of ICs and BCs for PDEs. Although we do not evaluate this phenomenon thoroughly, discussion on this issue can be found in the literature (e.g., Cuomo et al. 2022). Figure 3(c) shows the predicted velocity field where the ground truth is the constant horizontal velocity with no vertical velocity. This is an inverse modeling case where head and concentration data at collocation points are used to estimate a hydraulic conductivity field which is used for velocity calculation. As briefly mentioned, this is one advantage of PINNs, where inverse modeling can be performed in a Bayesian framework (see Karniadakis et al. 2021).

Table 1. Mean square error of three different cases with different decay coefficients and collocation points.

Parameter values	Evaluation points	Case 1	Case 2	Case 3
λ=0, N=500	Collocation points	6.09E-05	9.58E-05	5.62E-05
	Entire space & time	3.54E-03	2.61E-03	2.25E-03
λ=0.1, N=500	Collocation points	6.39E-05	9.16E-05	7.81E-05
	Entire space & time	3.40E-03	2.99E-03	1.62E-03
λ=0.5, N=500	Collocation points	9.76E-05	9.41E-05	7.46E-05
	Entire space & time	2.71E-03	2.06E-03	2.25E-03
λ=0.5, N=1,000	Collocation points	6.55E-04	9.76E-04	1.05E-03
	Entire space & time	4.81E-05	1.09E-04	1.08E-04

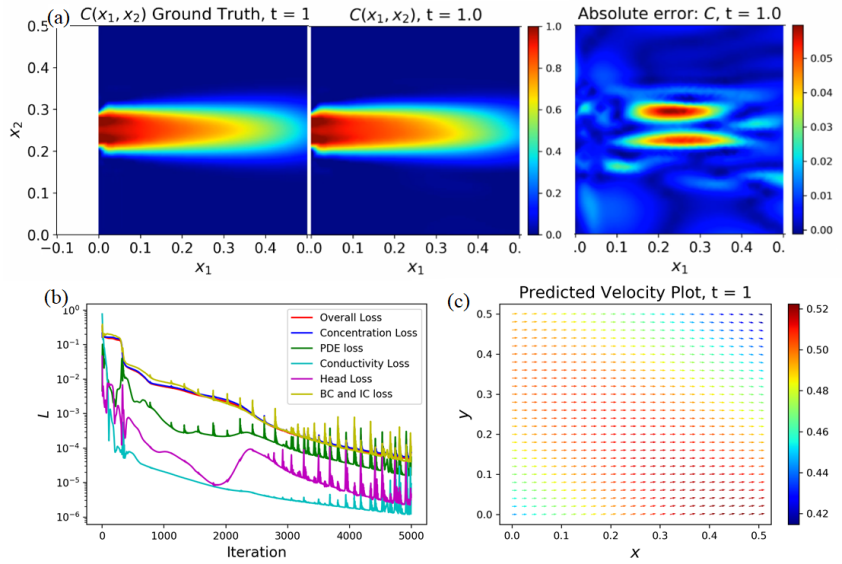

Figure 3. (a) Comparison of concentrations (analytical solution (left) and predicted one (middle)) in Case 3 (λ=0.5, N=500) at the last time step and (b) corresponding loss plots from different terms. (c) Predicted velocity field using inverse modeling with observed concentration and head for Case 3.

4 SUMMARY

We evaluate PINNs for multiple equations as both a forward and inverse modeler. Since data-driven DL is able to model large quantities of complex data to produce complex outputs without knowing underlying physical principles, physics-informed ML holds promise for subsurface applications. Although the analytical solution with constant parameter values of hydraulic conductivity, diffusion coefficient, and decay coefficient, the proposed framework can be readily applicable for more general systems of equations with a range of parameter space where analytical solution can't be used. In addition, PIML can improve both DL performance and the speed at which we train DL models, especially for incomplete, sparse, and/or noisy training data.

ACKNOWLEDGEMENTS

This work was supported by the Laboratory Directed Research and Development program at Sandia National Laboratories (Project # 229318). Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC (NTESS), a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration (DOE/NNSA) under contract DE-NA0003525. This written work is authored by an employee of NTESS. The employee, not NTESS, owns the right, title and interest in and to the written work and is responsible for its contents. Any subjective views or opinions that might be expressed in the written work do not necessarily represent the views of the U.S. Government. The publisher acknowledges that the U.S. Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this written work or allow others to do so, for U.S. Government purposes. The DOE will provide public access to results of federally sponsored research in accordance with the DOE Public Access Plan.

REFERENCES

Cai, S., Wang, Z., Wang, S., Perdikaris, P. and Karniadakis, G.E., 2021. Physics-informed neural networks for heat transfer problems. Journal of Heat Transfer, 143(6).

Cuomo, S., Di Cola, V. S., Giampaolo, F., Rozza, G., Raissi, M., & Piccialli, F. 2022. Scientific machine learning through physics–informed neural networks: where we are and what's next. Journal of Scientific Computing, 92(3), 88. https://doi.org/10.1007/s10915-022-01939-z

He, Q., Barajas-Solano, D., Tartakovsky, G., & Tartakovsky, A. M. 2020. Physics-informed neural networks for multiphysics data assimilation with application to subsurface transport. Advances in Water Resources, 141, 103610.

Kadeethum, T., O'Malley, D., Choi, Y., Viswanathan, H. S., Bouklas, N., & Yoon, H. 2022. Continuous conditional generative adversarial networks for data-driven solutions of poroelasticity with heterogeneous material properties. Computers & Geosciences, 167, 105212. DOI: 10.1016/j.cageo.2022.105212

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. 2021. Physics-informed machine learning. Nature Reviews Physics, 3(6), pp. 422-440. doi: 10.1038/s42254-021-00314-5.

Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., & Anandkumar, A. 2020. Fourier neural operator for parametric partial differential equations. arXiv preprint arXiv:2010.08895.

Lu, L., Jin, P., Pang, G., Zhang, Z., & Karniadakis, G. E. 2021. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. Nature machine intelligence, 3(3), pp. 218-229.

Paladino, O., Moranda, A., Massabò, M., & Robbins, G. A. 2018. Analytical Solutions of Three-Dimensional Contaminant Transport Models with Exponential Source Decay. Groundwater, 56(1), pp. 96-108.

Raissi, M., Perdikaris, P., & Karniadakis, G. E. 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational physics, 378, pp. 686-707.

Wang, S., Yu, X. and Perdikaris, P., 2022. When and why PINNs fail to train: A neural tangent kernel perspective. Journal of Computational Physics, 449, p.110768.