Finite Volume Neural Network: Modeling Subsurface Contaminant Transport

Timothy Praditia*[1], Matthias Karlbauer[2], Sebastian Otte[2], Sergey Oladyshkin[1], Martin V. Butz[2], Wolfgang Nowak[1]

[1]University of Stuttgart, [2]University of Tübingen

*Corresponding author

ICLR 2021 Workshop
Deep Learning for Simulation (simDL)

Is a Physics-Informed Neural Network enough for modeling spatio-temporal problems?

The short answer: **NO**!

The long answer:

- The network training depends heavily on the derivatives $\frac{\partial u}{\partial x}$ and $\frac{\partial u}{\partial t}$.
- When the training data is not distributed spatially or temporally, the approximation of derivatives deteriorates.
- Cannot generalize when tested against different boundary conditions.

Our proposed solution: **Hybrid model** combining the well-defined structure of the **Finite Volume Method** and learning ability of **Neural Ordinary Differential Equation**

The Model That Can Do It: Finite Volume Neural Network (FINN)

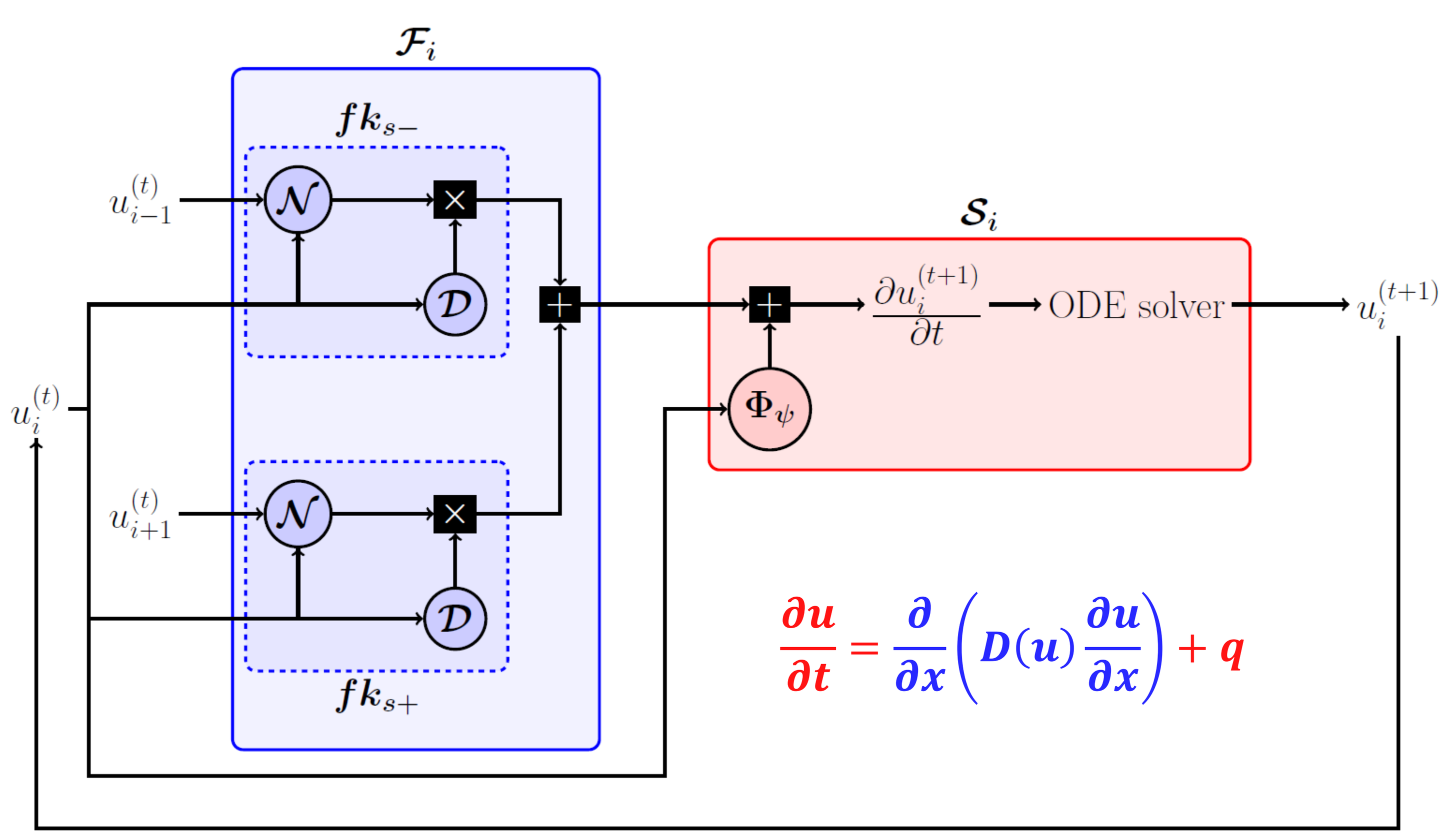


Illustration of the Flux and State Kernels in the FINN.

$\mathcal{F}$: Flux Kernel
(learn differentials, calculate fluxes and BCs, learn constitutive relationships)

$$\mathcal{F}_i = \Phi_\theta\left(u_{i-1}^{(t)}, u_i^{(t)}, u_{i+1}^{(t)}\right) = \sum fk_s \approx \oint_S \left(D(u)\frac{\partial u}{\partial x}.\hat{n}\right) ds$$

$\mathcal{S}$: State Kernel
(learn reaction term, integrate with ODE solver)

$$\mathcal{S}_i = \mathcal{F}_i + \Phi_\psi(u_i) \approx \frac{\partial u_i}{\partial t}$$

Experiment

Non-linear Diffusion-Sorption

Trichloroethylene (TCE) dissolved concentration:

$$\frac{\partial c}{\partial t} = \frac{D_e}{R}\frac{\partial^2 c}{\partial x^2}$$

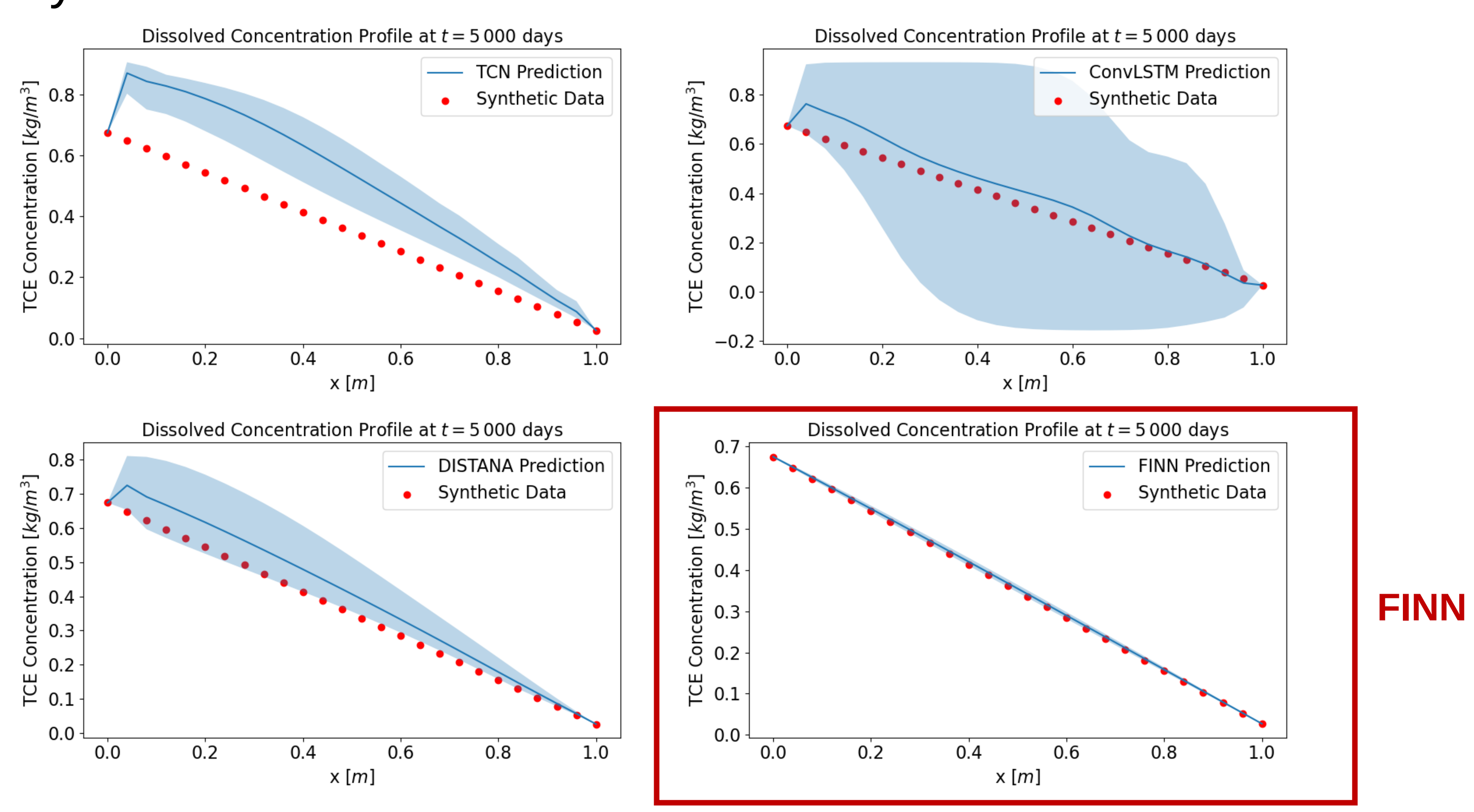
TCE total concentration:

$$\frac{\partial c_t}{\partial t} = D_e \phi \frac{\partial^2 c_t}{\partial x^2}$$

Dirichlet boundary condition at $x = 0$ and Cauchy boundary condition at $x = L$.

Results

Synthetic Dataset



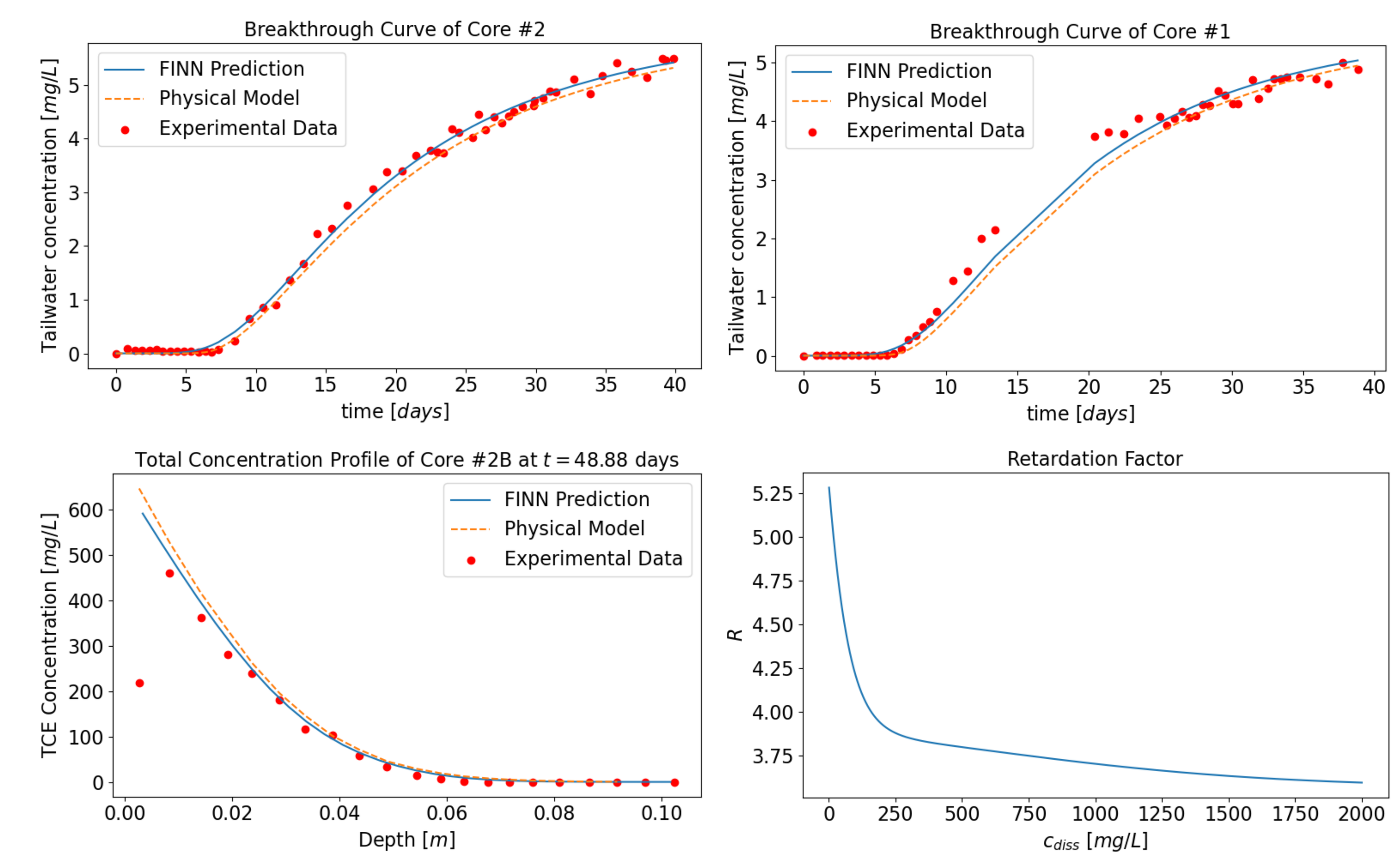
Dissolved concentration profile prediction mean (with confidence interval) at $t = 5\,000$ days compared with the test dataset obtained using TCN (top left), ConvLSTM (top right), DISTANA (bottom left), and FINN (bottom right).

The synthetic data is discretized into 26 control volumes and 2 000 time steps. We train using time steps 0 – 500 of the training dataset. The generalization is then tested by extrapolating the prediction for the training dataset until time step 2 000 and to predict a completely unseen test dataset (different boundary conditions).

Comparison of MSE values between different deep learning architectures

Method	Training	Extrapolated training	Test unseen	Parameters
TCN	$(7.9 \pm 5.4) \times 10^{-6}$	$(5.9 \pm 4.1) \times 10^{-3}$	$(3.0 \pm 1.2) \times 10^{-2}$	1 386
ConvLSTM	$(5.5 \pm 1.6) \times 10^{-6}$	$(4.9 \pm 5.7) \times 10^{-2}$	$(6.6 \pm 7.9) \times 10^{-2}$	1 496
DISTANA	$(1.9 \pm 1.1) \times 10^{-6}$	$(1.0 \pm 2.9) \times 10^{-2}$	$(1.6 \pm 4.0) \times 10^{-2}$	1 350
FINN	$(4.7 \pm 4.9) \times 10^{-5}$	$(1.1 \pm 1.2) \times 10^{-4}$	$(4.1 \pm 4.0) \times 10^{-5}$	528

Experimental Dataset



Dissolved concentration profile (at $x = L$) prediction of the FINN method (blue line) during training using data from core sample #2 (top left), during testing using data from core sample #1 (top right) and total concentration profile (at $t = t_{end}$) prediction using data from core sample #2B (bottom left). The predictions are compared with the experimental data (red circles) and the results obtained using the calibrated physical model (orange dashed line). The extracted retardation factor as a function of c is shown on the bottom right plot.

Conclusion and Future Work

- Using the numerical structure of the Finite Volume Method enables approximation of differential operators and conservative fluxes.
- FINN produces excellent generalization ability.
- Extension of FINN to applications with spatially heterogeneous soil parameters.
- Uncertainty quantification of the model using Bayesian method.

Acknowledgement

This work is funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2075 – 390740016 as well as EXC 2064 – 390727645. We acknowledge the support by the Stuttgart Center for Simulation Science (SimTech). Moreover, we thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting Matthias Karlbauer. Codes and data that are used for this paper can be found in the repository https://github.com/timothypraditia/finn.

References

R. T. Q. Chen, Y. Rubanova, J. Bettencourt and D. Duvenaud. Neural ordinary differential equations. In Advances in Neural Information Processing Systems, volume 31, 2018.

W. Nowak and A. Guthke. Entropy-based experimental design for optimal model discrimination in the geosciences. Entropy, 18(11), 2016.

C. Rackauckas, Y. Ma, J, Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, A. Ramadhan and A. Edelman. Universal differential equations for scientific machine learning. arXiv preprint, 2020.

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.