Efficient Uncertainty Quantification for Physics-Aware Machine Learning of Diffusion-Sorption Models

Universität Stuttgart SimTech

LS³

EGU26-3076

Stefania Scheurer[1,2], Riccardo Frenner[1,3], Tim Brünnette[1,2], Sergey Oladyshkin[1,2], Wolfgang Nowak[1,2]

[1] Institute for Modelling Hydraulic and Environmental Systems - Department of Stochastic Simulation and Safety Research for Hydrosystems, University of Stuttgart, Germany
[2] Stuttgart Center of Simulation Science, University of Stuttgart, Germany
[3] Institute of Industrial Automation and Software Engineering, University of Stuttgart, German

SCAN FOR FULL PAPER & CODE

SETTING

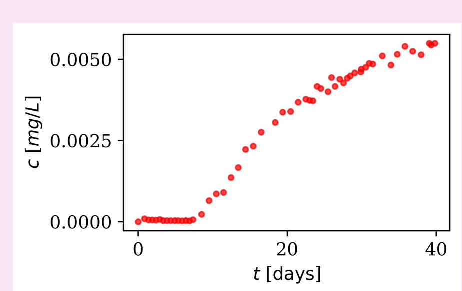
Given a physics-aware ML model

$$\frac{\partial c}{\partial t} = \mathcal{F}(c; \Theta(c) = NN_{\mathbf{w},\mathbf{b}}(c)) \quad t \in [0,T], x \in [0,L]$$

meaning a parameter is parameterized by a neural network,

$$\Theta(c) = NN_{\mathbf{w},\mathbf{b}}(c)$$

and trained with scarce observed data c^{obs}



using the FINite Volume Neural Network (FINN) Praditia et al. (2022)

FINN

- Replace to be learned $\Theta(c)$ of PDE with $NN_{\mathbf{w},\mathbf{b}}(c)$
- **Discretize** and solve PDE with FV scheme → **Extract** $\{c_i^{pred}(x^{obs}, t^{obs})\}_{i,j=1}^{N}$
- Optimize **loss** $\mathcal{L} = MSE(c^{obs}, c^{pred})$ via **backpropagation**

RESEARCH QUESTION

How can we

- with *indirect data*
- in a *high-dimensional, non-interpretable parameter space* $\{w, b\}$

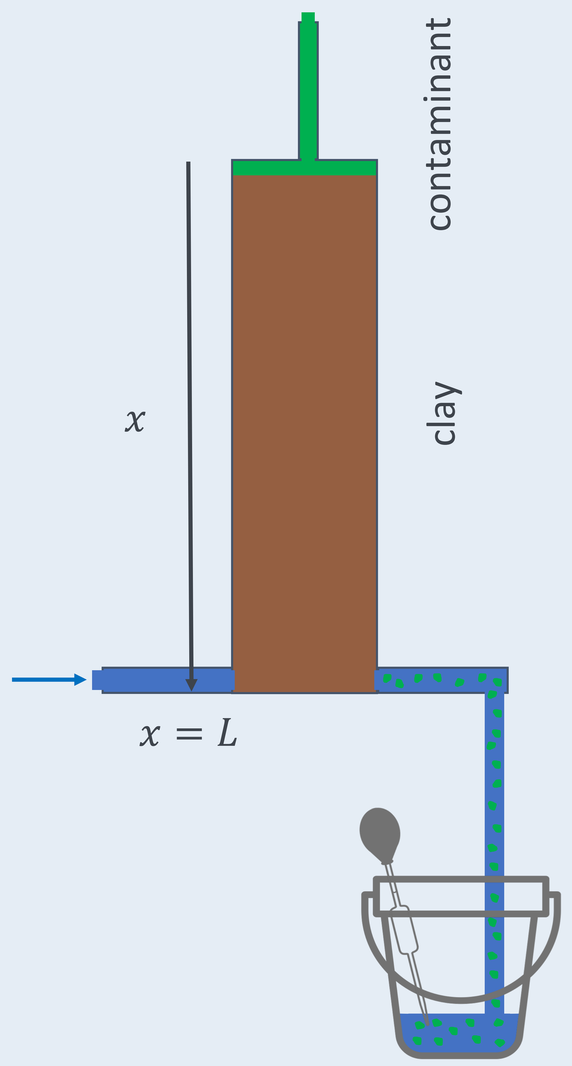
efficiently **quantify** the **uncertainty** for $\Theta(c) = NN_{w,b}(c)$?

METHOD

based on Liu et al., 2022.

Motivating Example

Experiment
Contaminant on top of clay, bottom end flushed with water. Contaminant concentration measured at different time points.



Mathematical description
Diffusion-sorption equation

$$\frac{\partial c}{\partial t} = \frac{D}{R(c)} \frac{\partial^2 c}{\partial x^2}$$

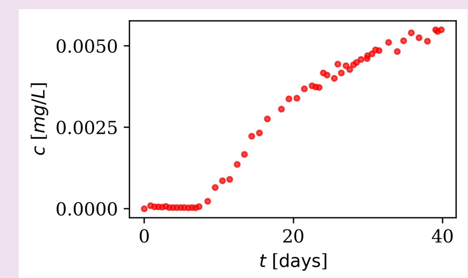
x: location
t: time
D: diffusion coefficient

and

to be learned parameter
$R(c)$: retardation factor

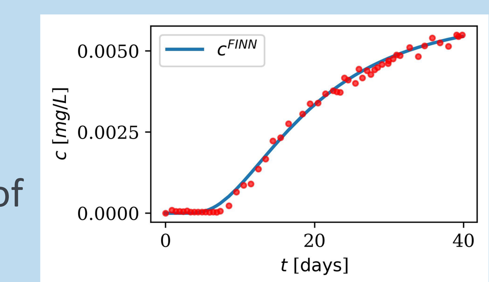
Step 1: Learn an Empirical Error Model

Given the **full dataset** $\{(x_i, t_i, c_i^{obs})\}_{i=1}^{N_{obs}}$ of the observable quantity (concentration c),



FINN

yields the **mean predictor** c^{FINN} and is shifted to be the median $\bar{c}$ of the dataset.



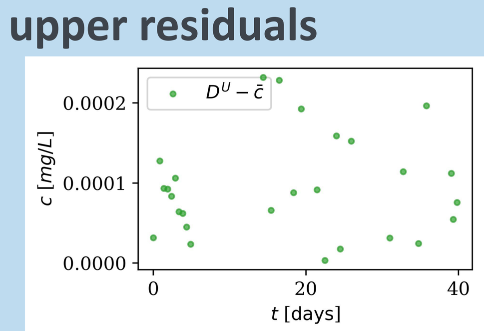
Splitting the full dataset into an **upper and lower dataset**

$$D^U = \left\{(x_i, t_i, c_i^{obs}) \, \mathbb{1}_{c_i^{obs} > \bar{c}(x_i, t_i)}\right\}_{i=1}^{N_{obs}},$$

$$D^L = \left\{(x_i, t_i, c_i^{obs}) \, \mathbb{1}_{c_i^{obs} < \bar{c}(x_i, t_i)}\right\}_{i=1}^{N_{obs}}$$

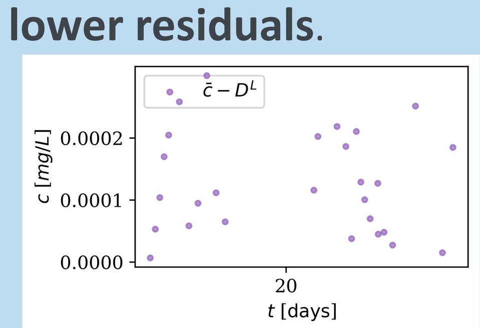
yields a dataset of

upper residuals



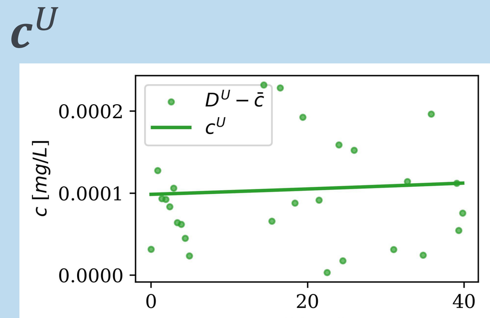
&

lower residuals.



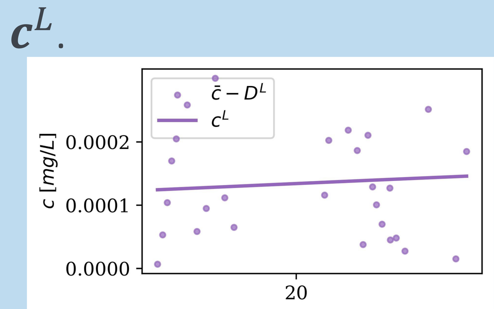
Training two feedforward NNs on the residual datasets yields an

upper residual predictor c^U



&

lower residual predictor c^L.

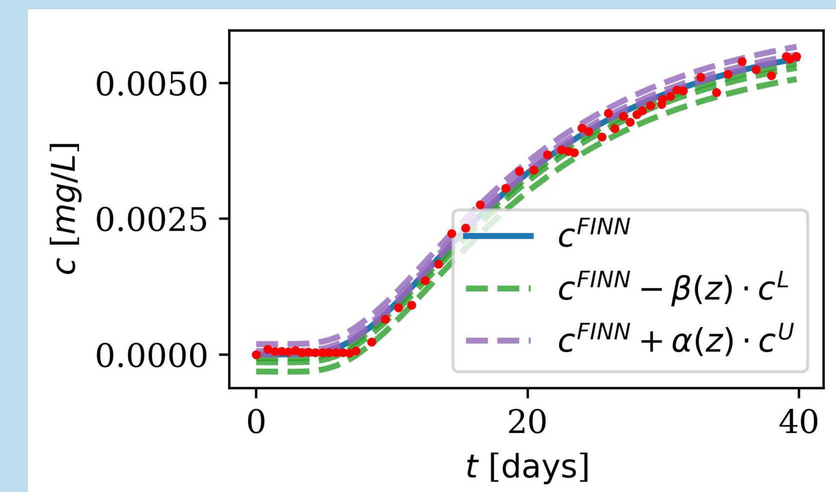


Step 2: Data-Driven Bootstrapping (DDB)

With $\mathbf{c}^{FINN}$, $\mathbf{c}^U$, and $\mathbf{c}^L$, we obtain a

$$\textbf{quantile function } Q(z) = \begin{cases} \mathbf{c}^{FINN} + \alpha(z) \cdot \mathbf{c}^U & z \in [0.5, 1] \\ \mathbf{c}^{FINN} + \beta(z) \cdot \mathbf{c}^L & z \in [0, 0.5) \end{cases} = F^{-1}_{\mathcal{C}^l | \mathcal{C} = c^{obs}}(z)$$

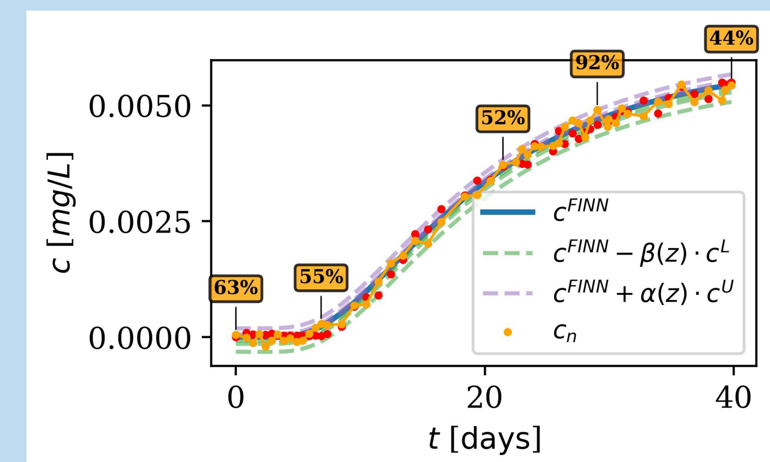
with parameters $\alpha(z)$ and $\beta(z)$ to be optimized according to wanted quantile z.



Sampling a **quantile level** $z_{n,t_i} \sim \mathrm{Unif}(0,1)$ for each time step t_i yields

the n-th **bootstrap sample**

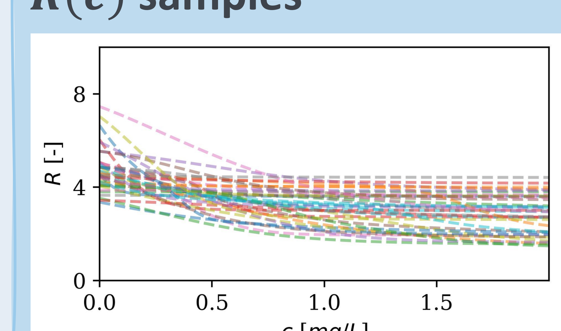
$$c_n = \{Q(z_{n,t_i})\}_{i=1}^{N_{obs}}$$



and with **FINN** the corresponding **retardation factor bootstrap sample** r_n.

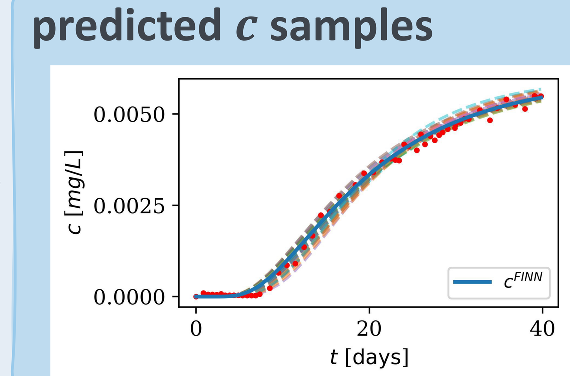
Repeating this N^{boot} times yields

$R(c)$ samples



&

predicted c samples

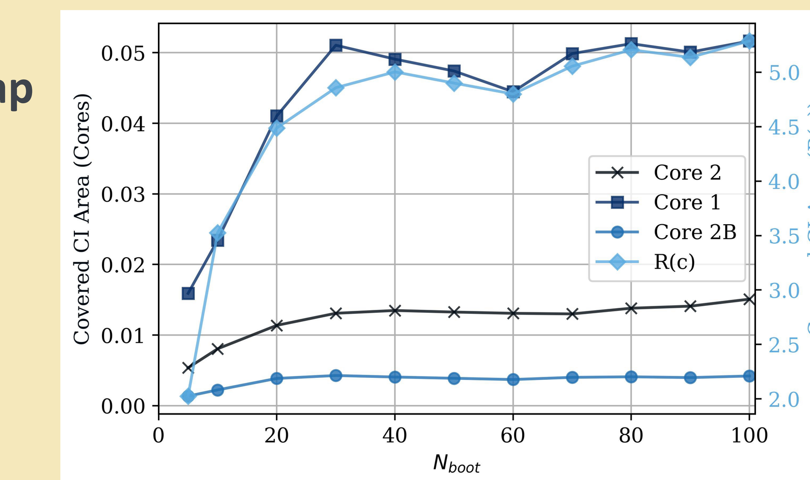


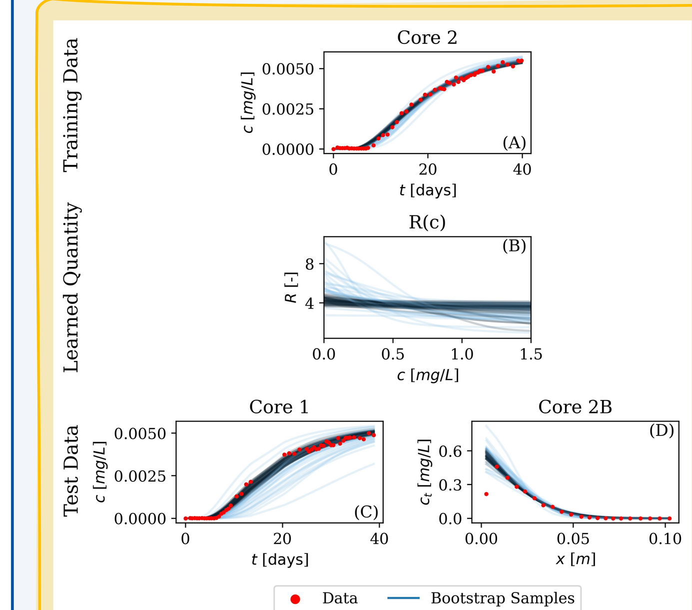
for confidence interval construction.

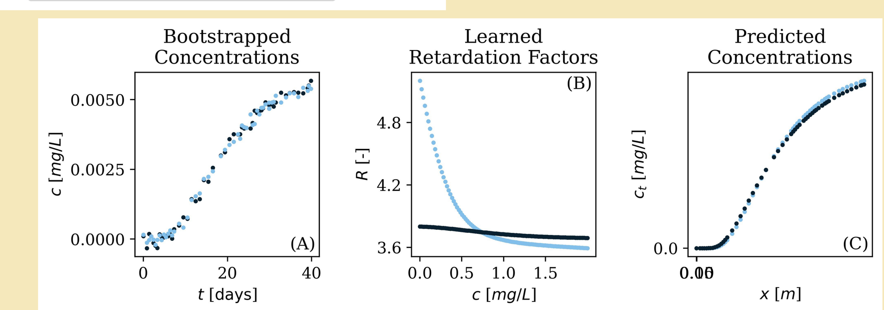
RESULTS

Results for training data (Core 2) and test data (Core 1, Core 2B).

70 bootstrap samples achieve converged confidence interval.

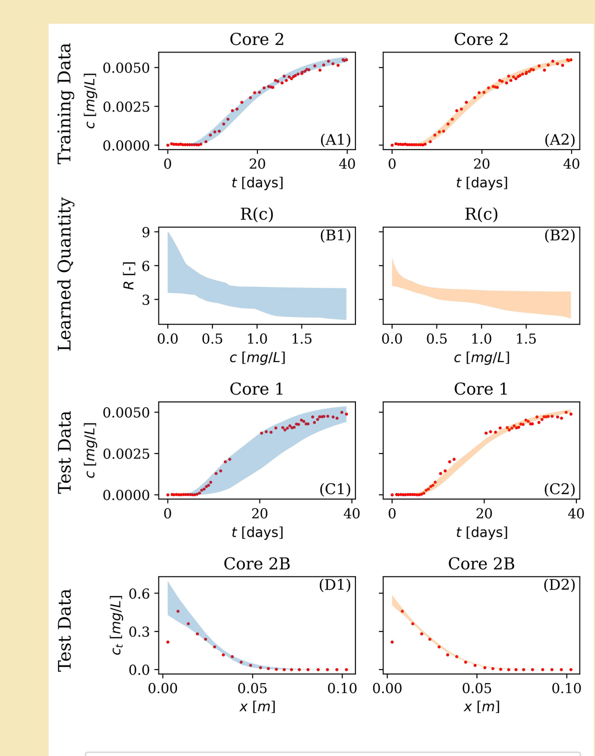




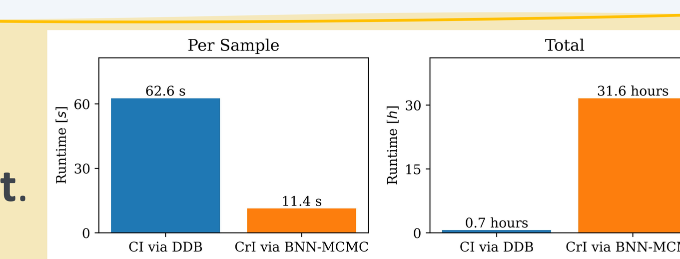


DDB **identifies apparent local optima** of the physics-constrained learning problem for a **reliable confidence interval.**

DDB better captures increased uncertainty by broader intervals for test data, whereas credible intervals via MCMC remain narrow due to the **influence of the prior**.



DDB is more **efficient**.



References
Praditia, T., Karlbauer, M., Otte, S., Oladyshkin, S., Butz, M. V., and Nowak, W. (2022). Learning groundwater contaminant diffusion-sorption processes with a finite volume neural network. Water Resour. Res. 58:e2022WR033149. doi: 10.1029/2022WR033149.
Liu, S., Zhang, P., Lu, D., and Zhang, G. (2022). PI3NN: Out-of-Distribution-Aware Prediction Intervals from Three Neural Networks. ICLR 2022 Poster.
Scheurer, S., Frenner, R., Brünnette, T., Oladyshkin, S., and Nowak, W. (2026). Efficient Confidence Interval Computation for Physics-Aware Machine Learning of Diffusion-Sorption Models. Frontiers in Water. Accepted for publication.

EGU Outstanding Student & PhD candidate Presentation contest