University of Stuttgart

Poster ID 12561 EGU 2025

Surrogate-assisted Bayesian inference with ERT data for contaminant transport modeling in the subsurface

Maria Fernanda Morales Oreamuno[1], Nino Menzel[2], Florian Wagner[2], Sergey Oladyshkin[1], Wolfgang Nowak[1]

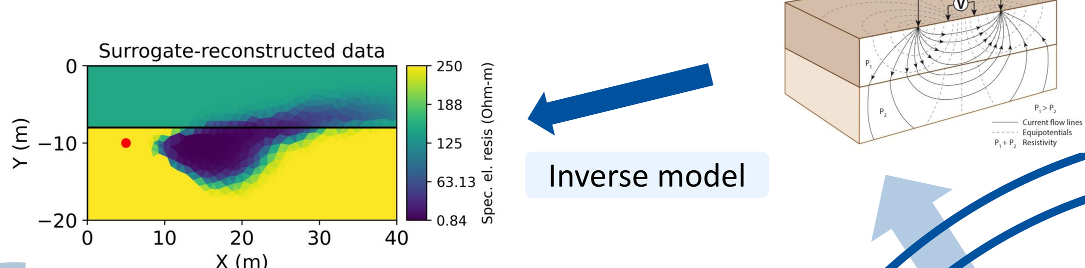
[1] Institute for Modelling Hydraulic and Environmental Systems - Department of Stochastic Simulation and Safety Research for Hydrosystems, University of Stuttgart, Stuttgart, Germany

[2] Geophysical Imaging and Monitoring (GIM), RWTH Aachen University, North Rhine-Westphalia, Germany

Electrical Resistivity Tomography (ERT) Data

✓ Non-invasive geophysical tool

✓ Can provide insights into subsurface properties and fluid movement→ sensitive to T and P changes



We use ERT measurements to do inference

– High dimensional measurement data

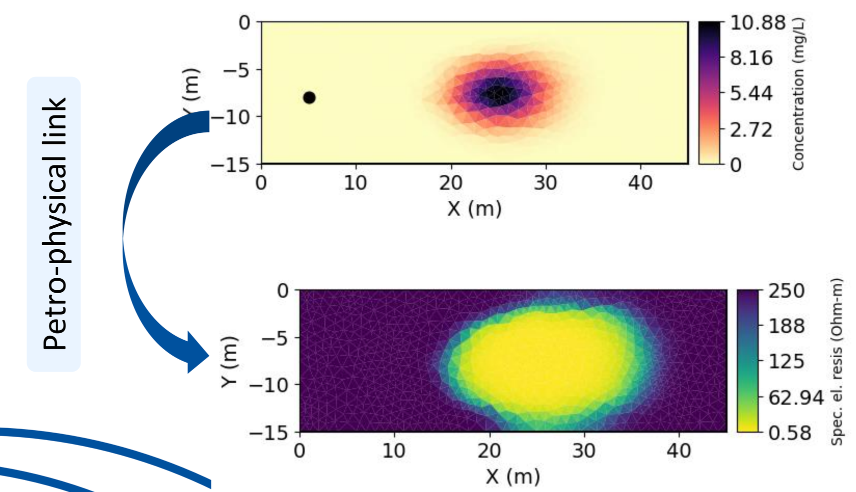
Objective

Use **ERT** measurements to reduce predictive uncertainty for **contaminant transport simulations**

Challenges

- Uncertainties in subsurface and contaminant-related parameters
- Computationally-expensive simulations
  - ERT data is high dimensional, which is a challenge for likelihood estimations and surrogate training

Contaminant Transport Model



2D non-reactive transport

- Homogeneous test case
- We consider 4 uncertain parameters (ω)
  - K
  - Φ
  - α_L
  - Contaminant strength (s_r)

We map from concentrations to **electrical resistivity data** through a petro-physical link

Surrogate Model

Surrogate models $\mathcal{S}$ are used to approximate a forward model's outputs (y), given variable inputs ($\boldsymbol{\omega}$), <u>at a fraction of the time</u>.

$$y \equiv \hat{y} = \mathcal{S}(\boldsymbol{\omega}, \theta)$$

Noisy GPR[1] approach

$$\hat{y} = \mathcal{S}(\omega, c) \approx N(\mu^*(\cdot), K(\cdot,\cdot))$$

PCE[2] approach with Bayesian solver

$$\hat{y} = \mathcal{S}(\omega, c) \approx \sum_{i=0}^{P} c_i \psi(\omega)$$

We build surrogates to estimate electrical resistivity data for different time steps of interest

Bayesian Inference

To quantify uncertainty in contaminant transport predictions by updating prior parameter beliefs using ERT data ($\boldsymbol{y}_o$)

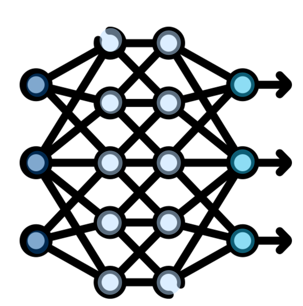
$$p(\omega|\boldsymbol{y}_o) = \frac{p(\boldsymbol{y}_o|\omega)p(\omega)}{p(\boldsymbol{y}_o)}$$

Likelihood free inference methods

✓ Avoids intractable likelihood computations in high dimensions

✓ Our approach: Simulation-based inference[3]

✓ Surrogate-assisted approach: to generate training samples



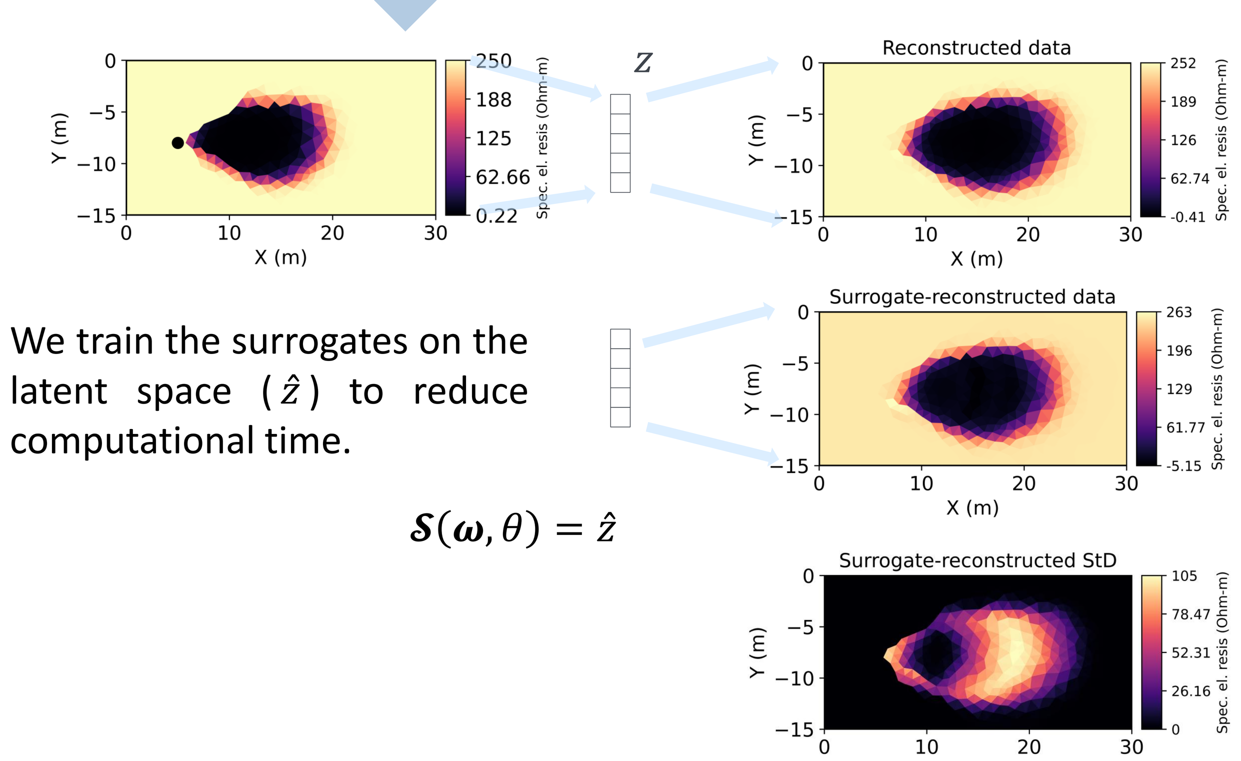
Neural density operator

Research questions:

- How much does **ERT** data help reduce uncertainty in contaminant transport predictions?
- What are the impacts of dimensionality reduction on the accuracy of **surrogate-assisted Bayesian inference**?

Dimension Reduction Approaches

We represent our quantity of interest in a low-dimensional latent space (z).



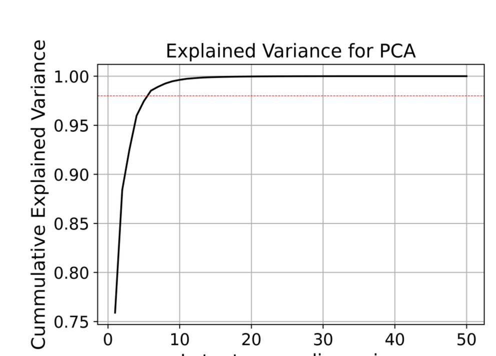
We train the surrogates on the latent space ($\hat{z}$) to reduce computational time.

$$\mathcal{S}(\omega, \theta) = \hat{z}$$

We consider **PCA** :

✓ Linear transformation

– Global encoder: Smoothens data

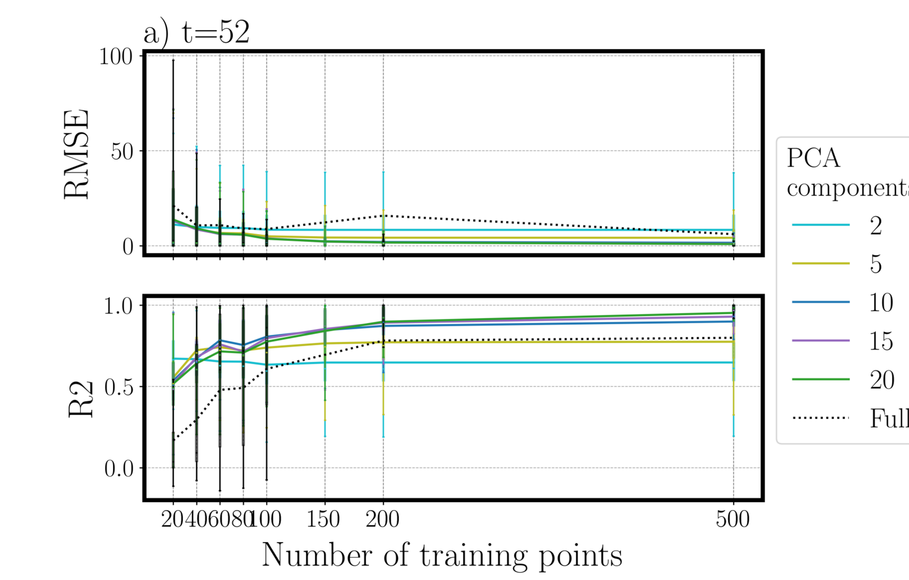


Results

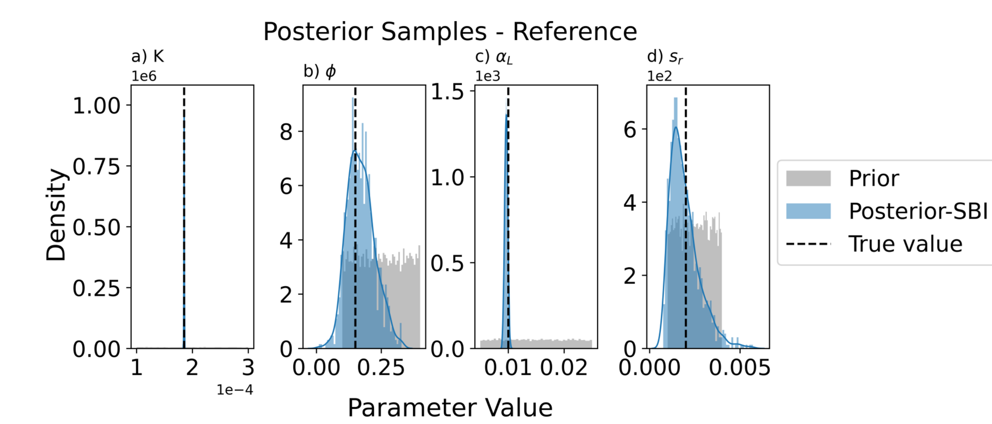
Predictive accuracy

✓ PCA-based surrogates show similar or higher predictive accuracy as per-cell surrogate approach (full)

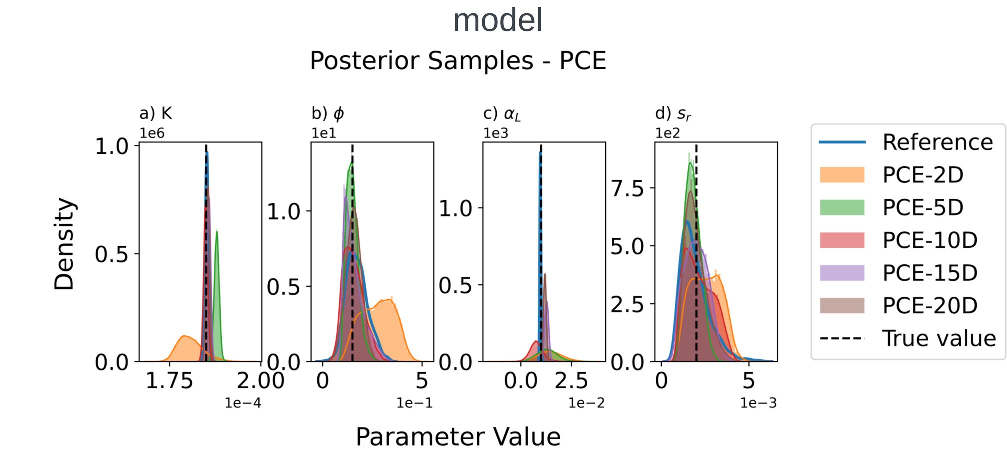
✓ PCA-based surrogates exhibited slightly higher predictive uncertainty, but overall good generalization.



Validation criteria for PCA-PCE surrogate models for different latent space size ($\hat{z}$)



SBI-obtained posterior samples using the reference model



SBI-obtained posterior samples using PCA-PCE surrogates for training

Bayesian Inference

- SBI using 10,000 model runs successfully recovered posterior distributions close to the synthetic truth.
  - Computationally intensive due to large number of model runs needed
- PCA-trained surrogates enabled a **faster** SBI training
- Using ≥98% retained variance yielded posterior convergence comparable to full-dimensional reference models.
- ERT data provides insights into field-scale parameters (e.g., permeability, porosity, dispersivity)

Outlook

Implement non-linear dimension reduction techniques to reduce surrogate-related uncertainty.

Implement methodology to optimal experimental design simulations for continuous monitoring.

Sources

[1] Williams, C. K., & Rasmussen. Gaussian processes for ma-chine learning, Volume 2. MIT press Cambridge, MA, 2006.

[2] Oladyshkin, S. and Nowak, W.. Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion. Reliability Engineering and System Safety, 106:179–190, 2012. ISSN 0951-8320.

[3] Greenberg, D., Nonnenmacher, M., & Macke, J. (2019, May). Automatic posterior transformation for likelihood-free inference. In International Conference on Machine Learning (pp. 2404-2414). PMLR.

[4] Tejero-Cantero et al., (2020). sbi: A toolkit for simulation-based inference. Journal of Open Source Software, 5(52), 2505, https://doi.org/10.21105/joss.02505

SimTech

BGE BUNDESGESELLSCHAFT FÜR ENDLAGERUNG

BVRox