

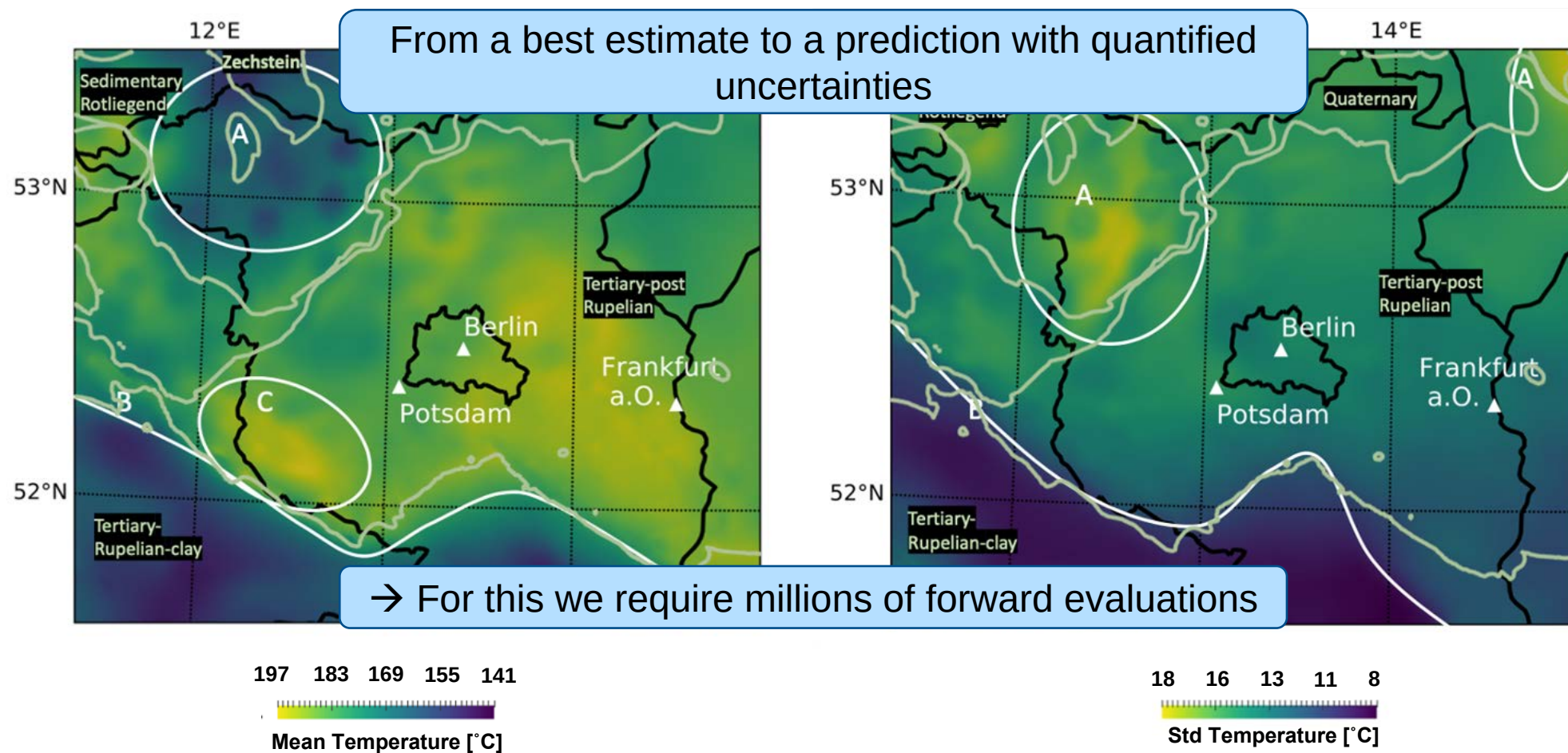
INTERPRETABLE PHYSICS-BASED MACHINE LEARNING MODELS FOR MULTI-QUERY APPLICATIONS

Denise Degen^{1,2}, Yulia Gruzdeva², Nicolas Hayek², Marthe Faber², Cristian Siegel²,
Mauro Cacace²

¹ Institute of Applied Geosciences, TU Darmstadt, Schnittspahnstr. 9, 64287 Darmstadt

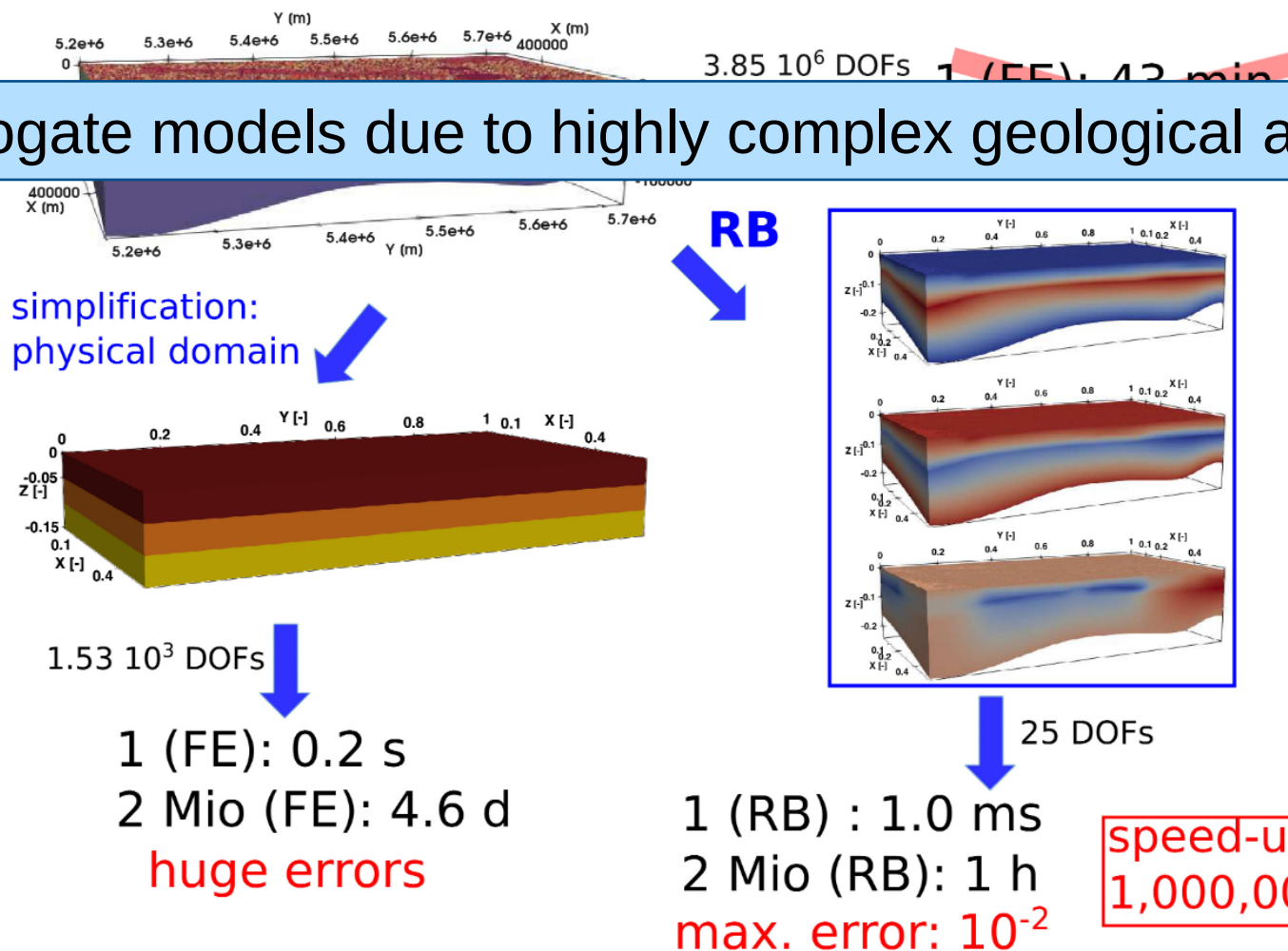
² GFZ Helmholtz Centre for Geosciences, Telegrafenberg, 14473 Potsdam

MOTIVATION



WHY SURROGATES

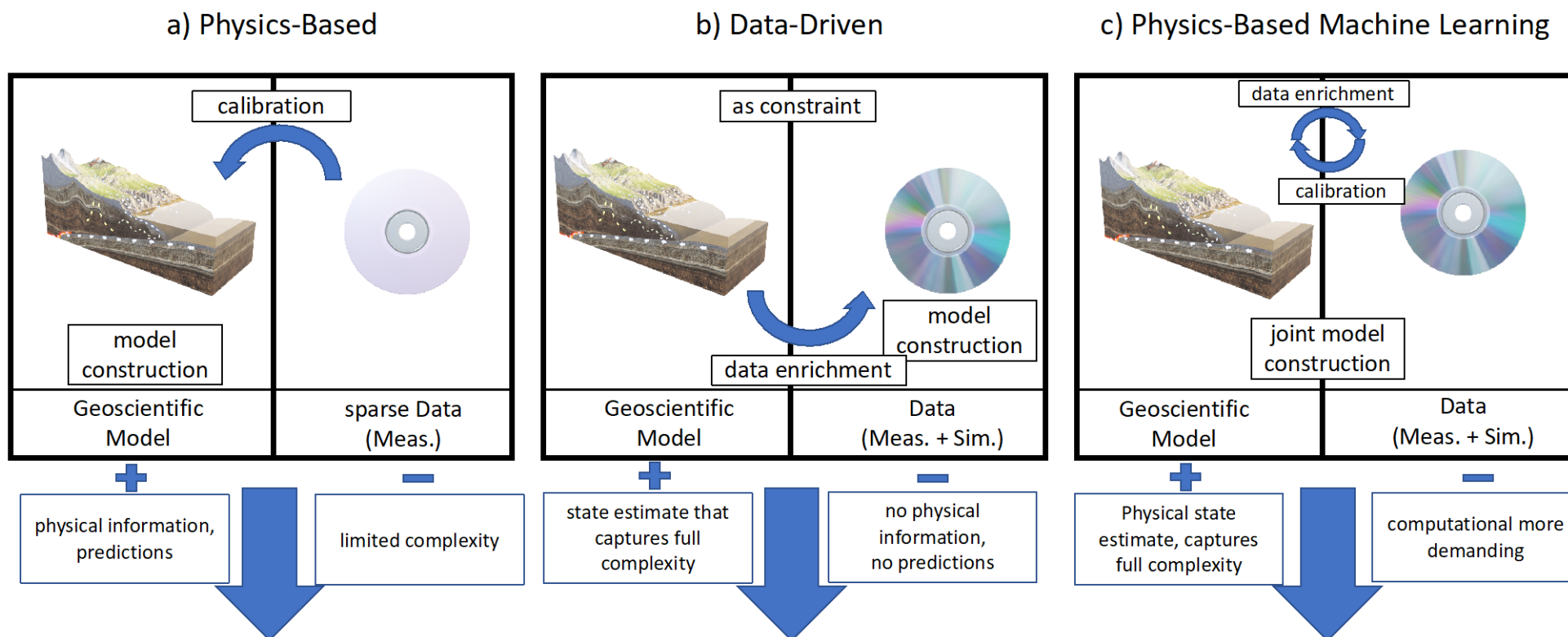
Need for surrogate models due to highly complex geological and physical models



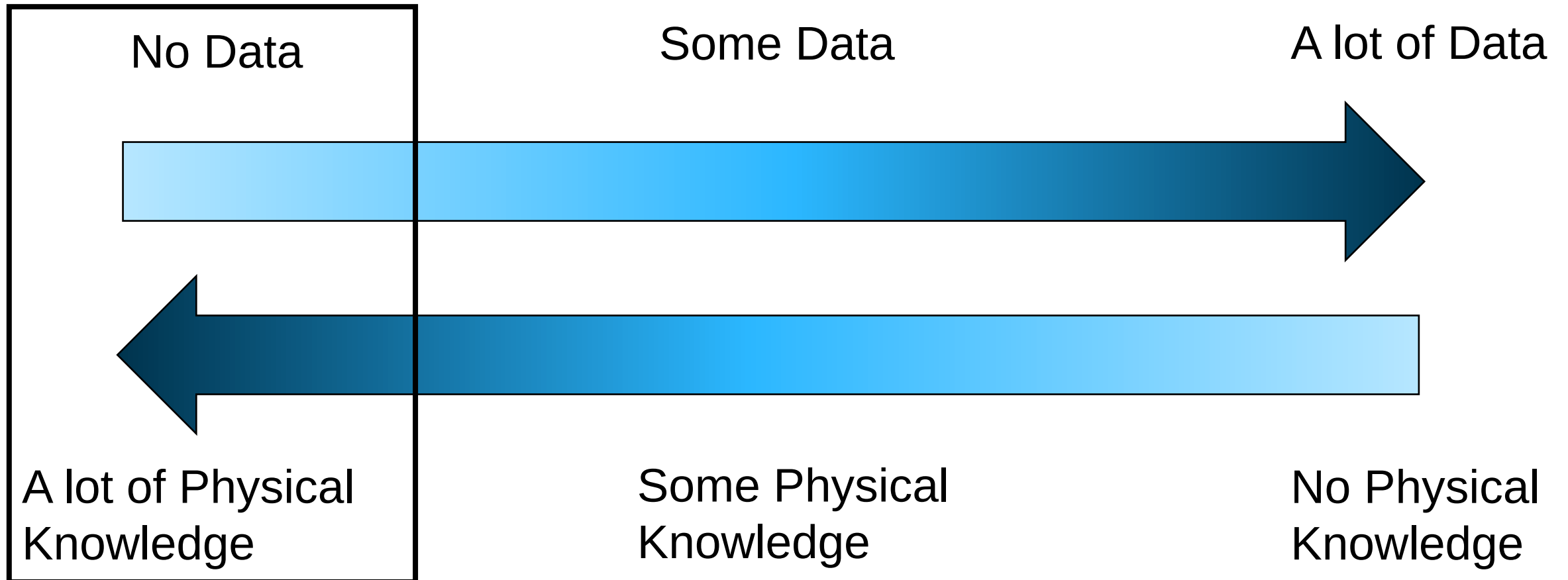
speed-up:
1,000,000

Degen et al. (2020a)

SURROGATE MODEL TECHNIQUES



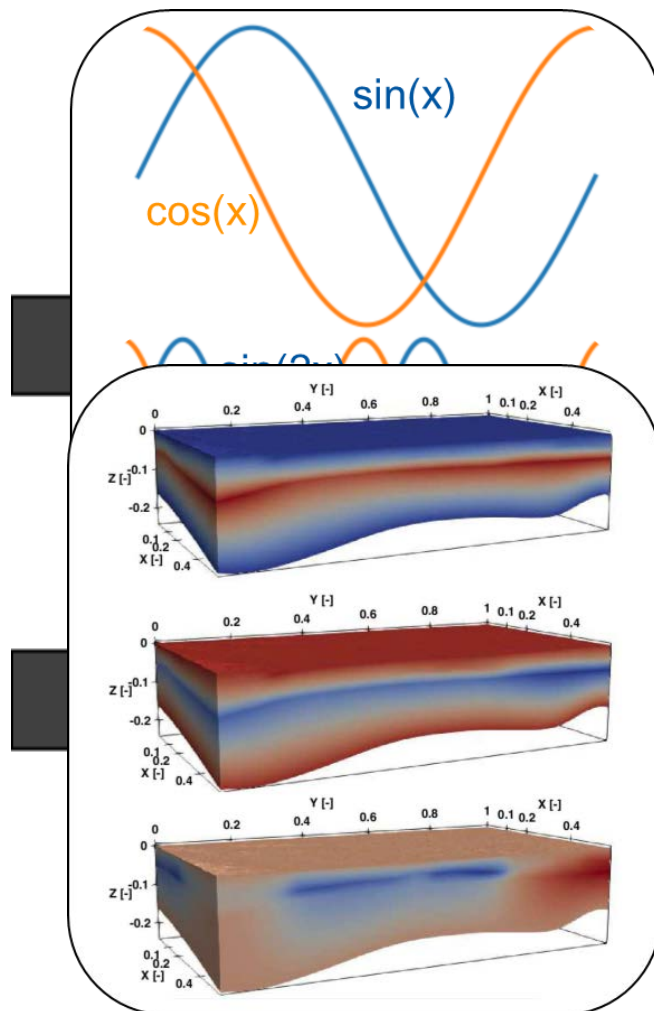
DATA VS. PHYSICS



NON-INTRUSIVE REDUCED BASIS METHOD

**Time
Domain**

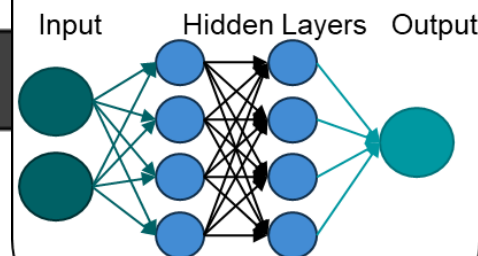
**Full Order
Model**



**Frequency
Domain**

**Reduced
Order Model**

e.g., (Deep) Neural Network

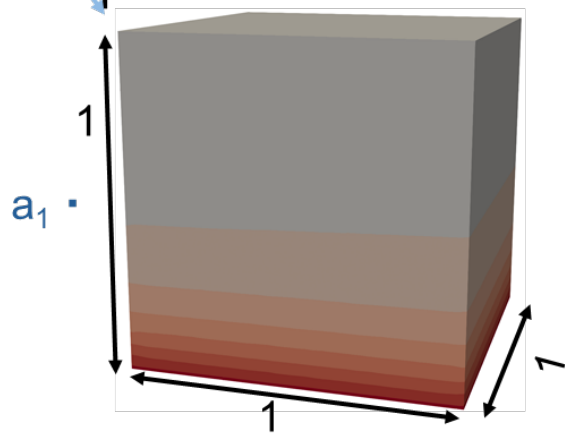
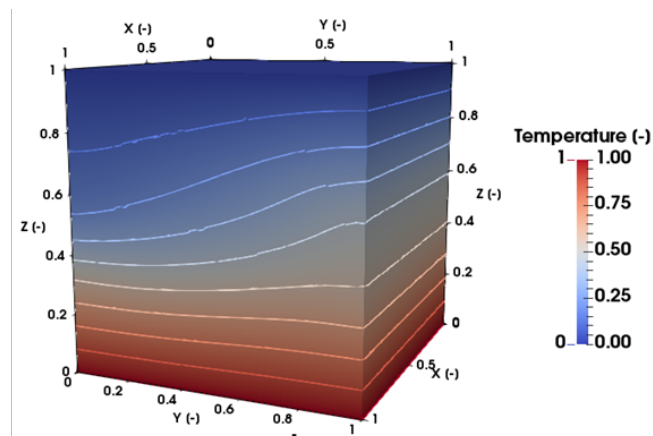


CONCEPTUAL EXAMPLE

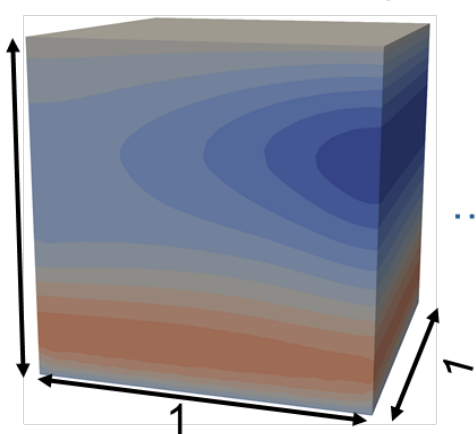
Surrogate Model of
the Temperature

Diffusive Term
("low frequency
information")

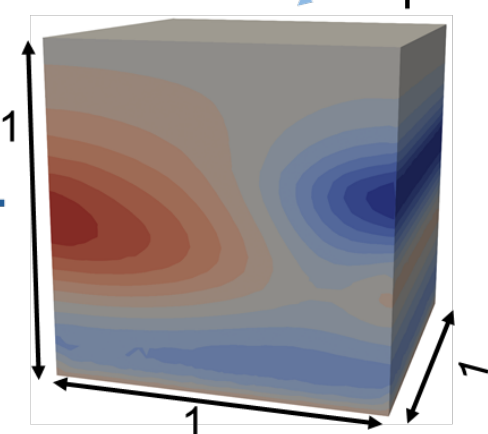
Advective Terms
(high frequency
information")



+ a₂ ·



... + a_n ·

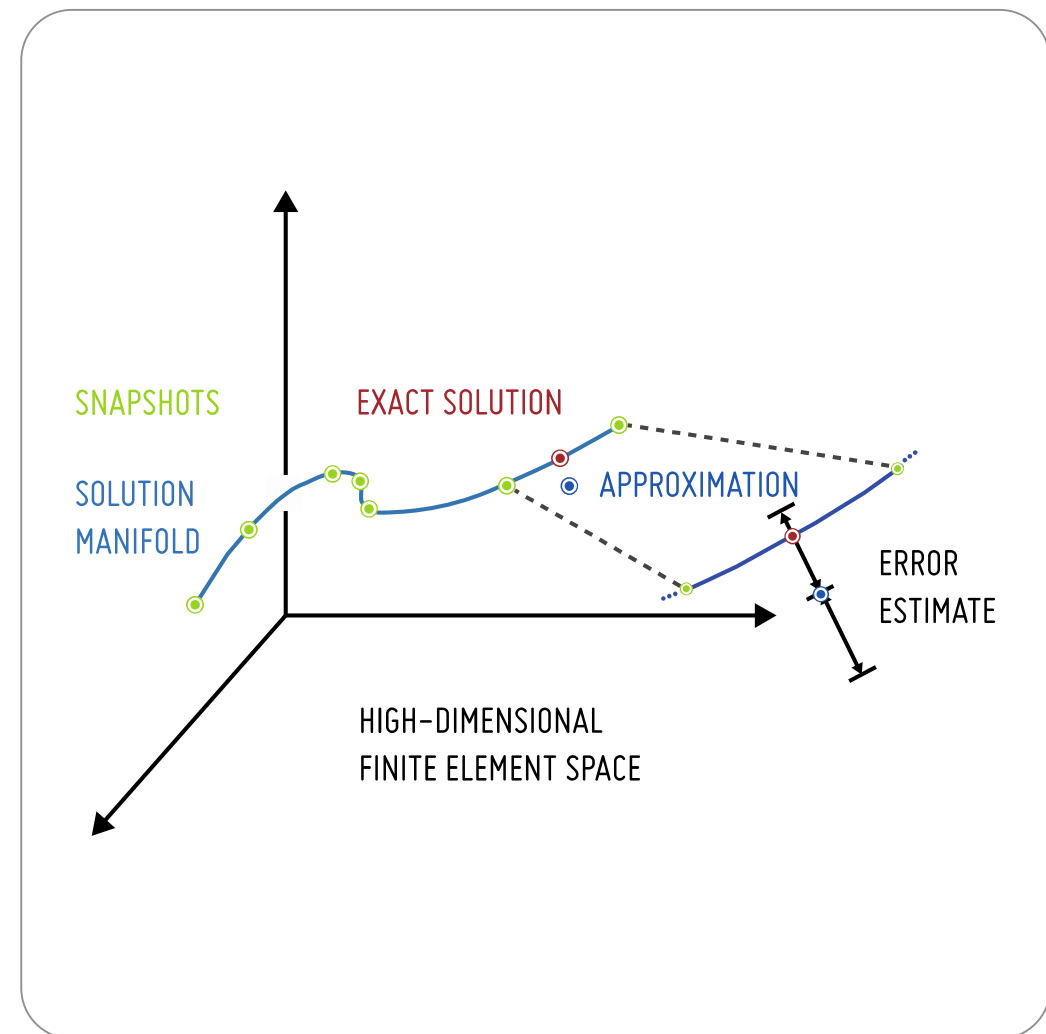


PROCEDURE

1. Construct a training and test data set
 - Latin Hypercube sampling for training set
 - Random sampling for test set
2. Scale the data-sets
→ training and test snapshots
3. Perform the POD with pre-defined tolerance

$$\frac{\sum_{i=1}^r \lambda_i^2}{\sum_{i=1}^N \lambda_i^2} \leq \varepsilon$$
 - Input: training snapshots
 - Output: basis functions
4. Perform the projection (machine learning)
 - Input: training parameters (e.g., material properties)
 - Output: reduced coefficients

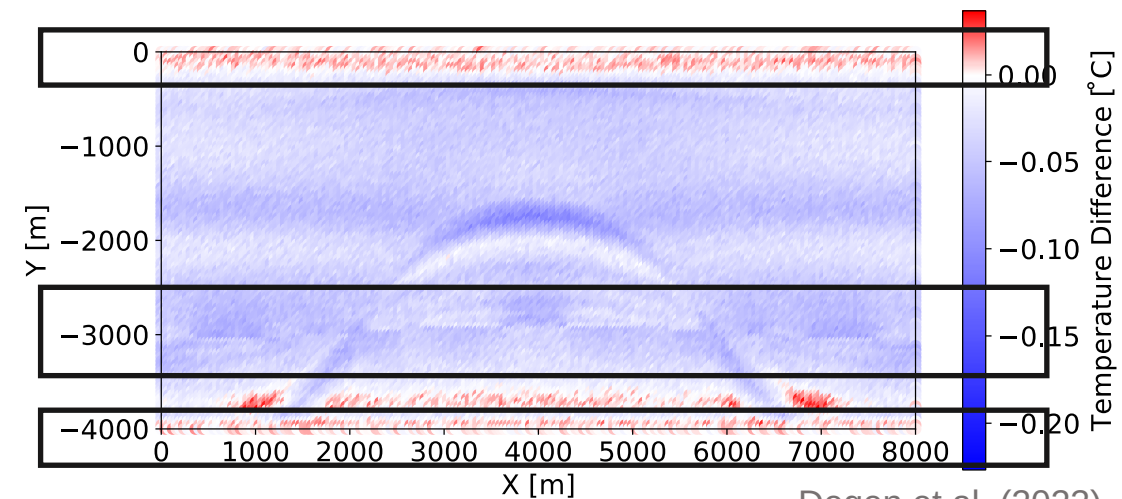
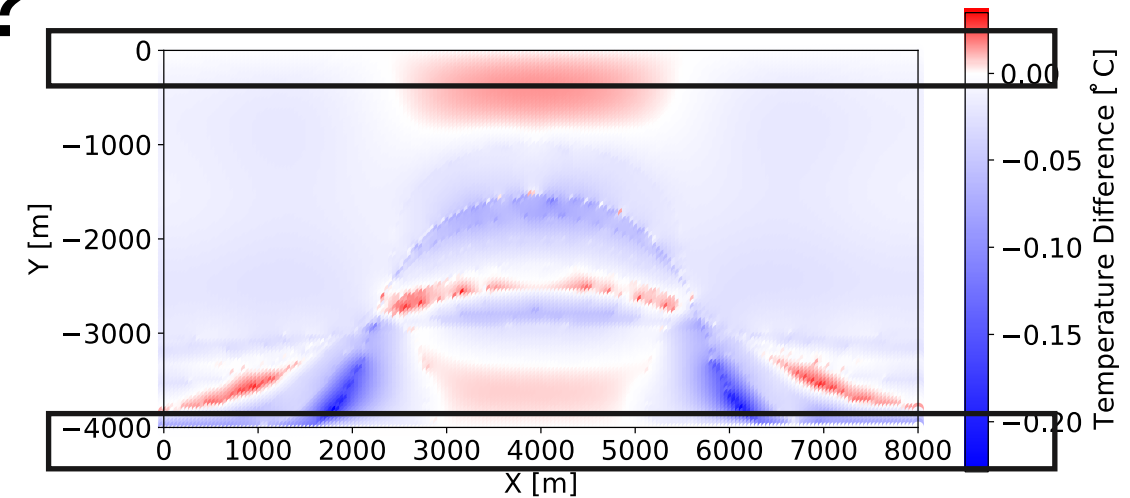
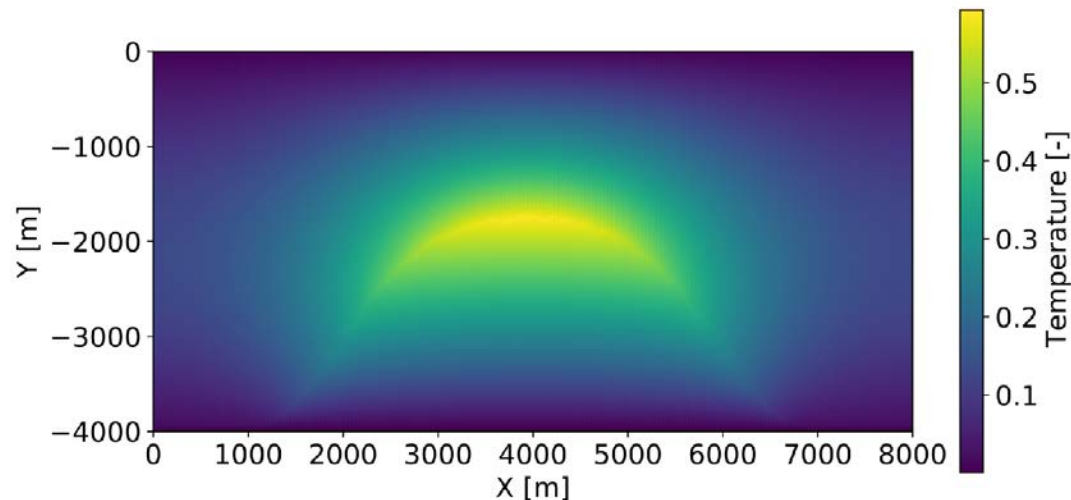
5. Surrogate: $u_{rb}(\theta) = \sum_{i=1}^r \beta_{rb}^{(i)}(\theta) \psi_i \in V_{rb}$



Degen et al. (2020b)

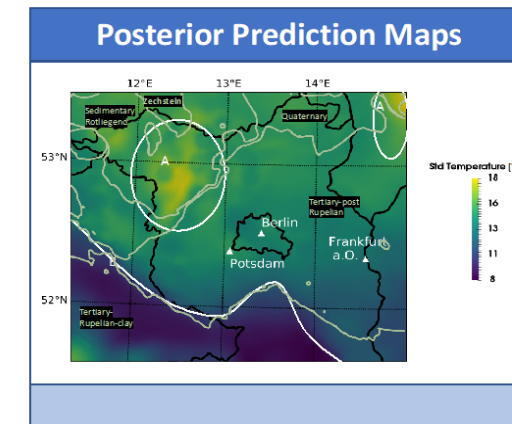
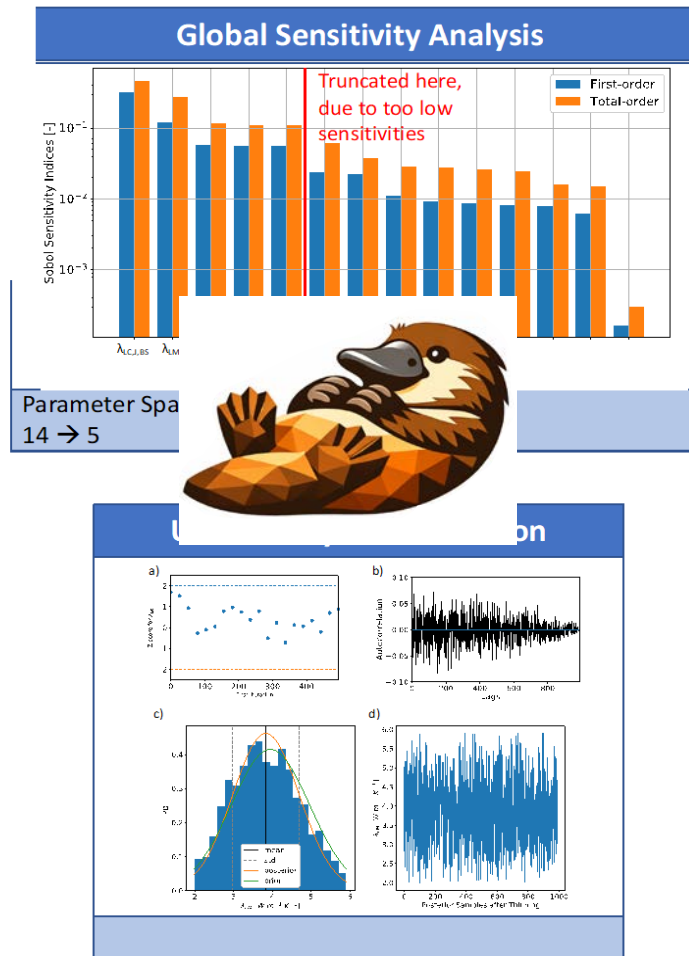
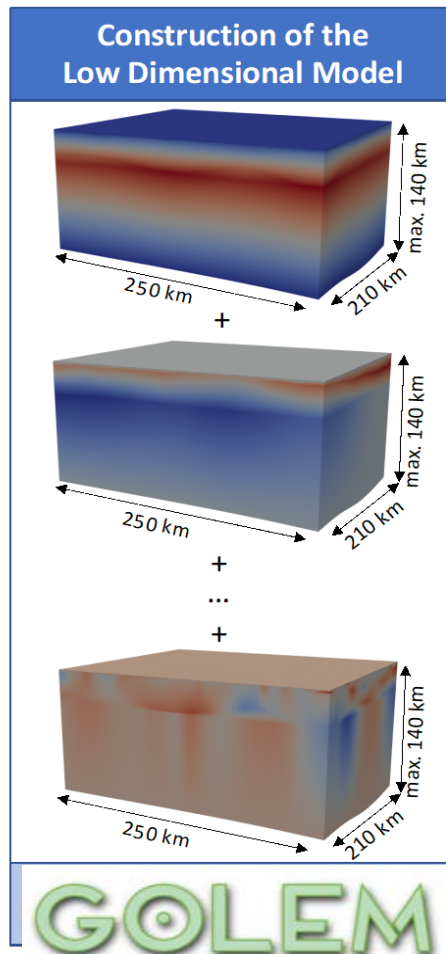
WHY PHYSICS-BASED MACHINE LEARNING?

- Reduction of the amount of data required for training the model
- Error distributions following physical laws
- Reduced training times



Degen et al. (2023)

WORKFLOWS



Degen et al. (2022a)

Degen & Wellmann (2026), Sobol
(2001), Wrainwright et al. (2014)

SENSITIVITY ANALYSES

A sensitivity analysis determines the **influence** of the model parameters on the **model response**.

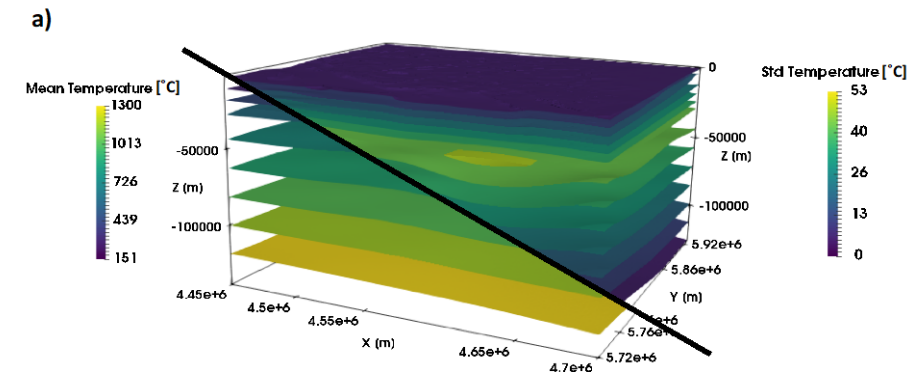
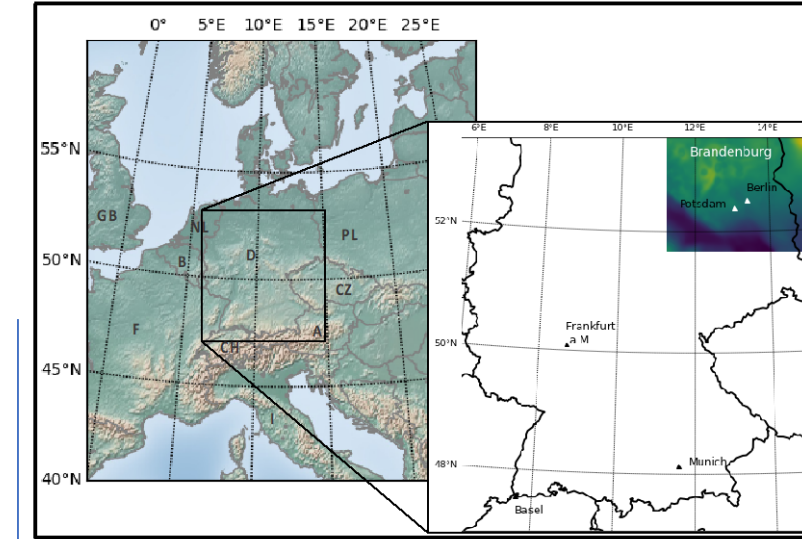
- Two categories:
 - Local sensitivity analysis
 - Global sensitivity analysis
- Local sensitivity analysis:
 - Local influence → with respect to pre-defined reference
 - Vicinity of the input parameters
 - No correlations considered
 - Very few evaluations
- Global sensitivity analysis:
 - i.e. Sobol sensitivity analysis → variance-based
 - Parameter distribution does not need to be known a priori
 - Possible consideration of correlations
 - Many evaluations

BERLIN-BRANDENBURG: UNCERTAINTY QUANTIFICATION

- Markov chain Monte Carlo
- Aims at reducing the uncertainty by incorporating data
- Based on Bayes Theorem

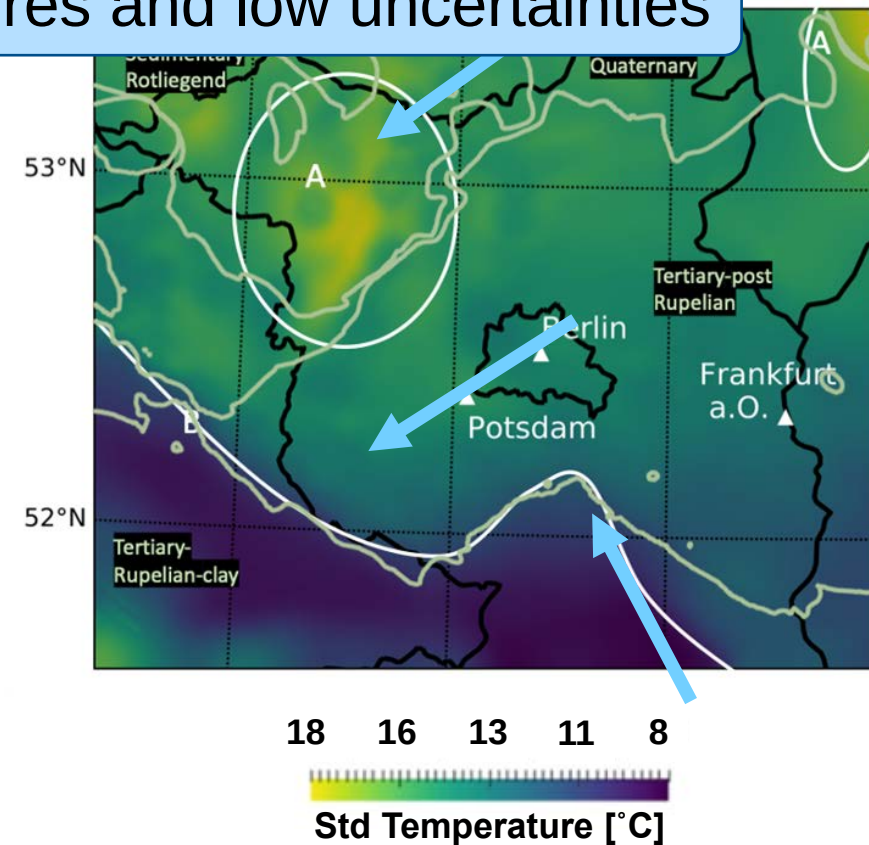
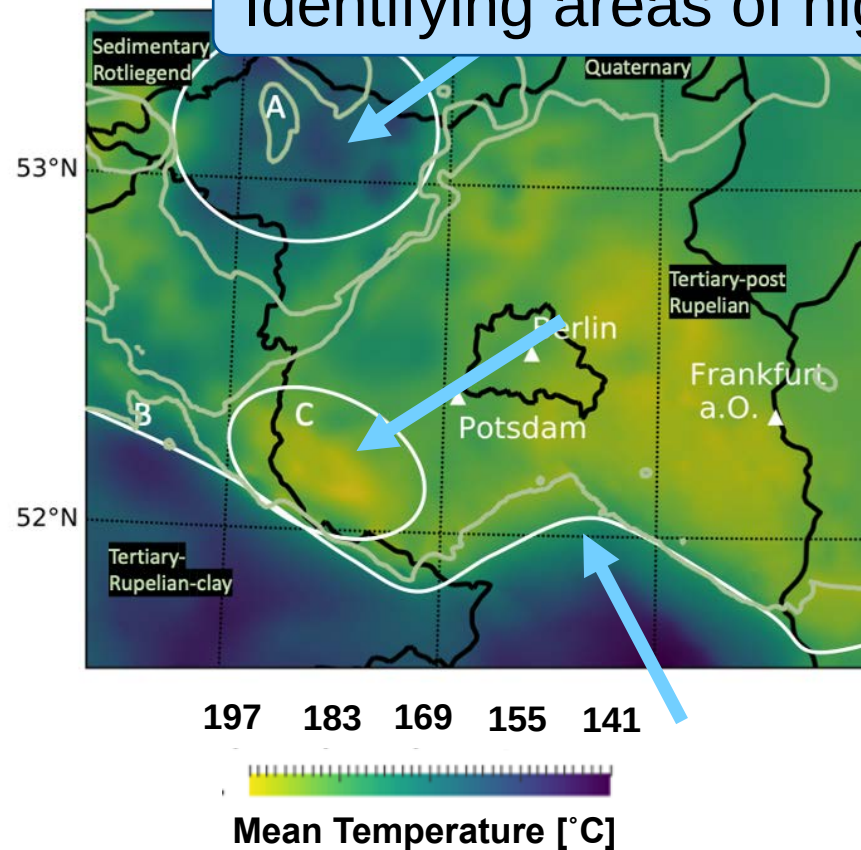
$$P(u \mid y) \propto P(y \mid u)P(u)$$

- $P(u \mid y)$: posterior
- $P(y \mid u)$: likelihood
- $P(u)$: prior



BERLIN-BRANDENBURG: UNCERTAINTY QUANTIFICATION

Identifying areas of high temperatures and low uncertainties



Degen et al. (2022a)

RELIABLE AND INTERPRETABLE MODELS

- Governing Equations (coupled THM simulation):

$$\frac{1}{M_b} \frac{\partial p_f}{\partial t} + \nabla \cdot q_D = 0,$$

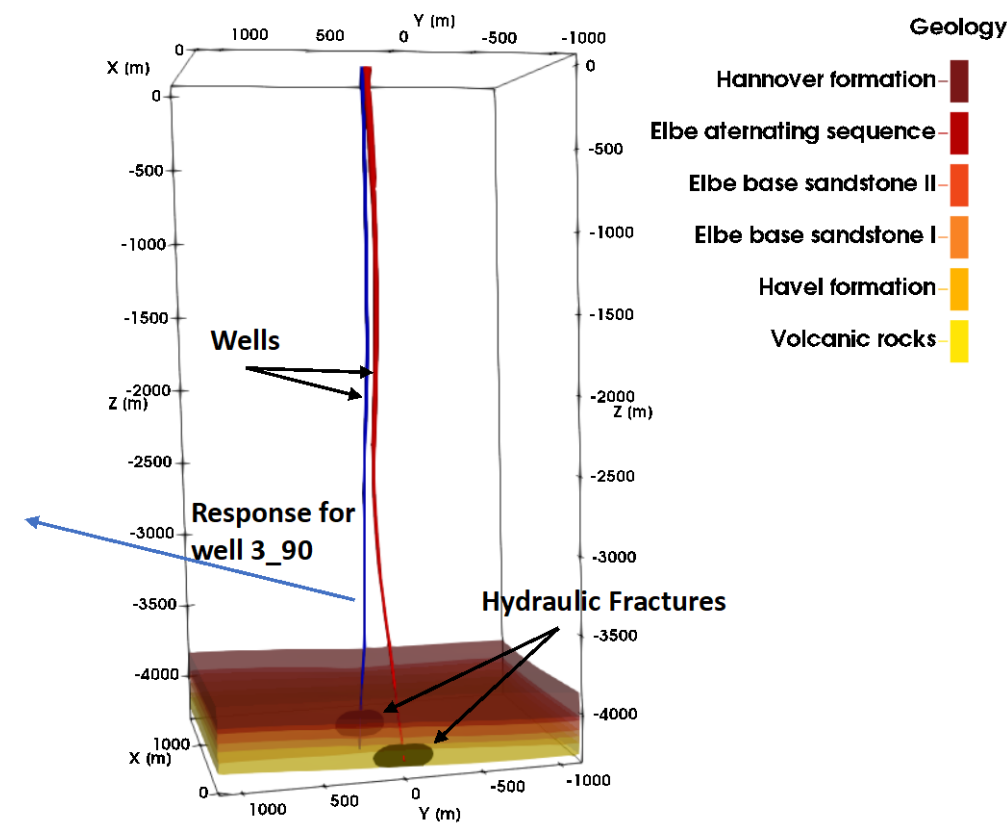
$$\text{with } q_D = -\frac{k}{\mu_f} (\nabla p_f - \rho_f g),$$

$$(\rho c)_b \frac{\partial T}{\partial t} + \nabla \cdot ((\rho c)_f q_D T - \lambda_b \nabla T) = 0,$$

$$\nabla \cdot (\sigma' - \beta p_f I) + \rho_b g = 0.$$

- Simulations performed with Golem (Cacace & Jacquy, 2017)

b)

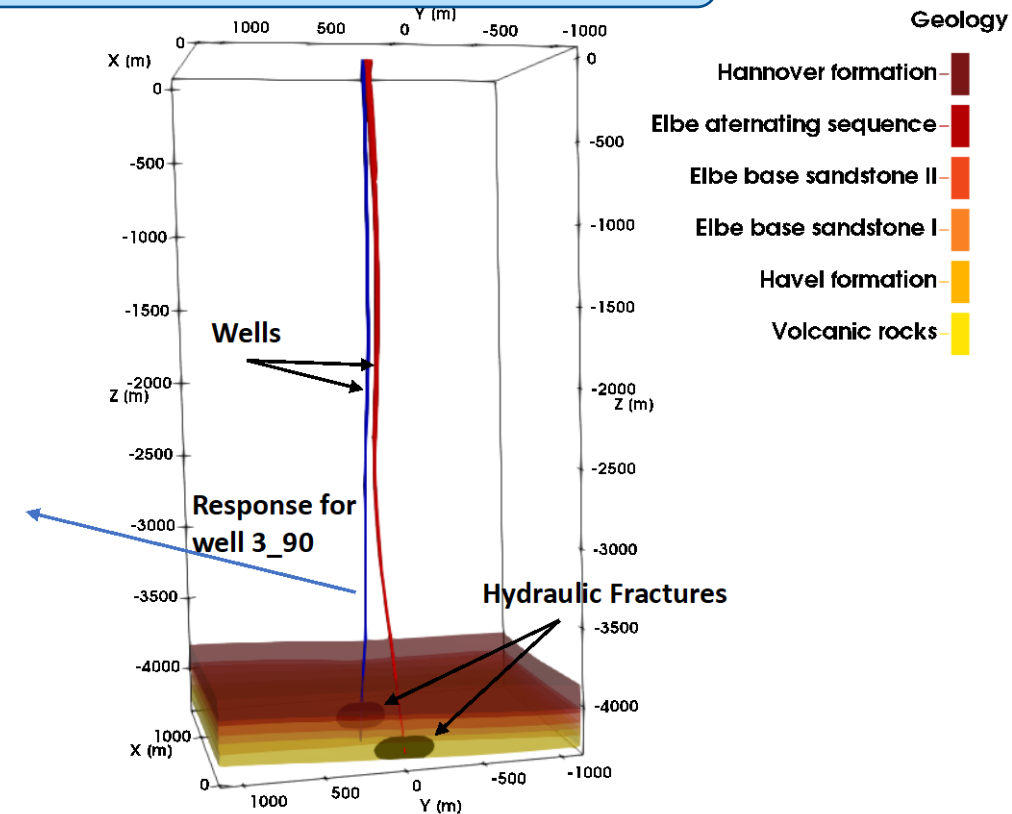
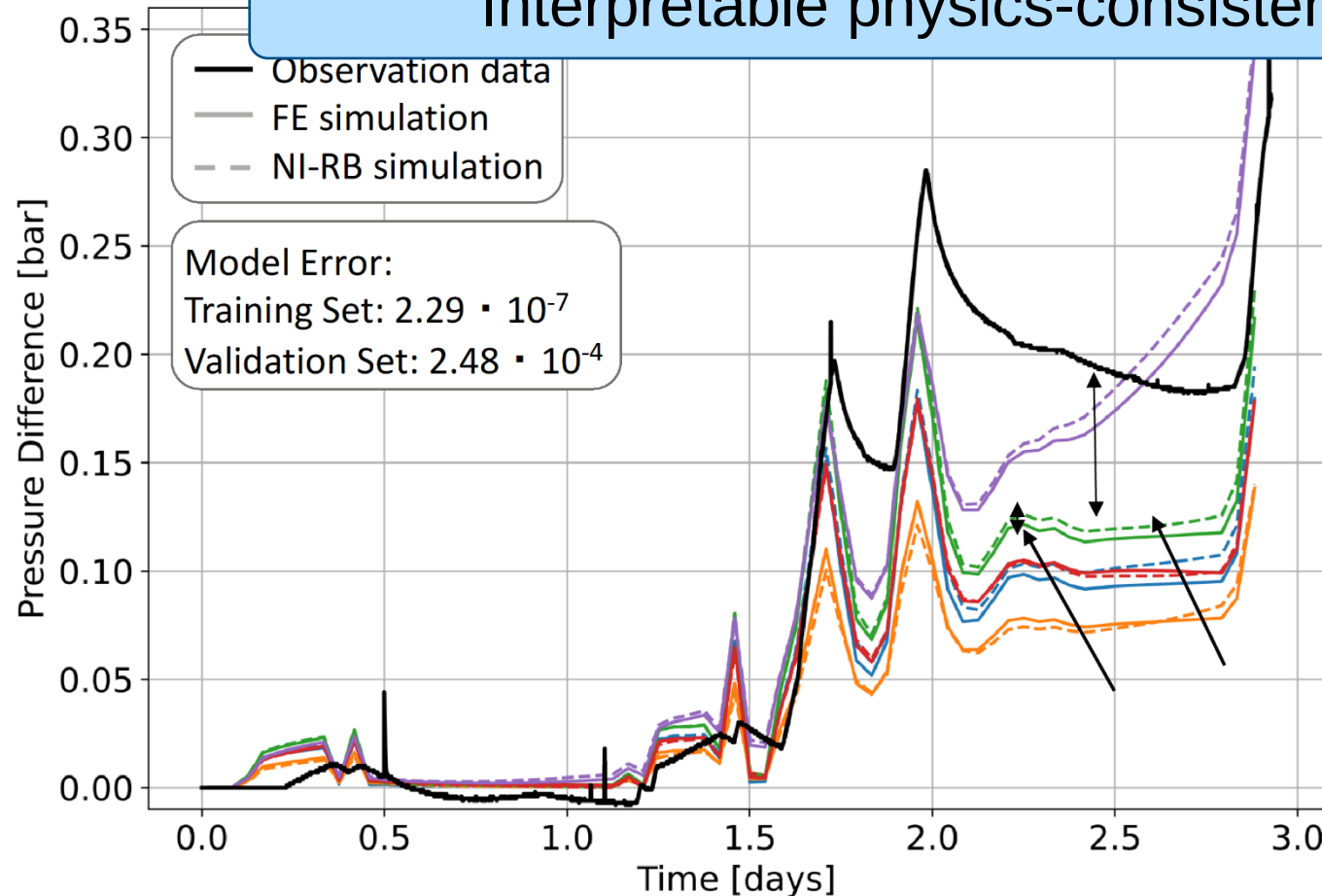


Degen et al. (2022b)

RELIABLE AND INTERPRETABLE MODELS

a)

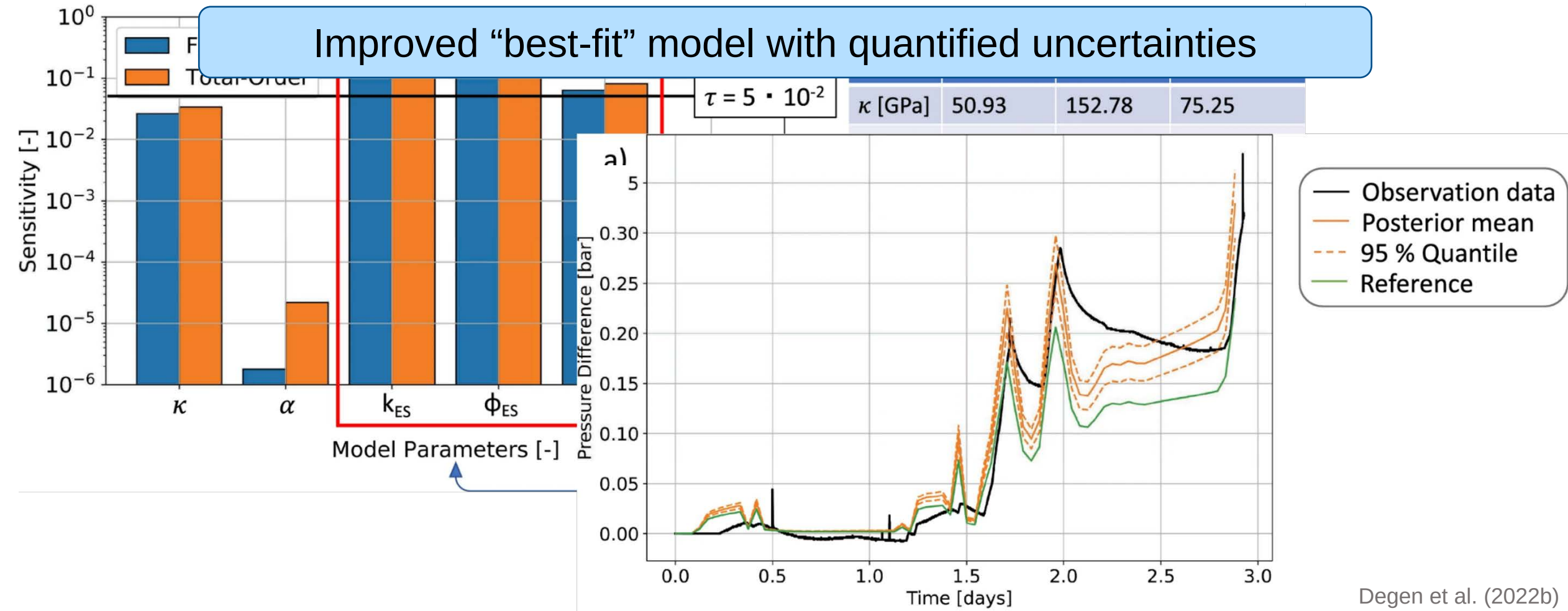
Interpretable physics-consistent surrogate models



Degen et al. (2022b)

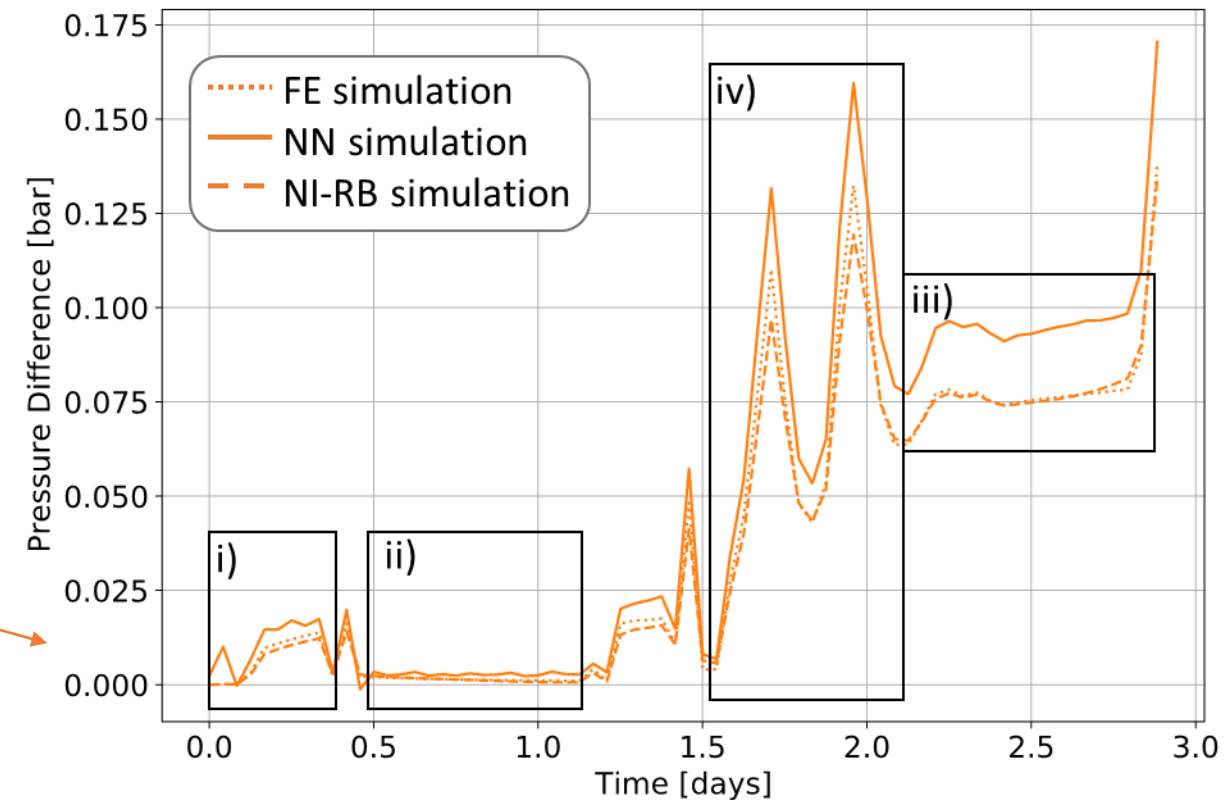
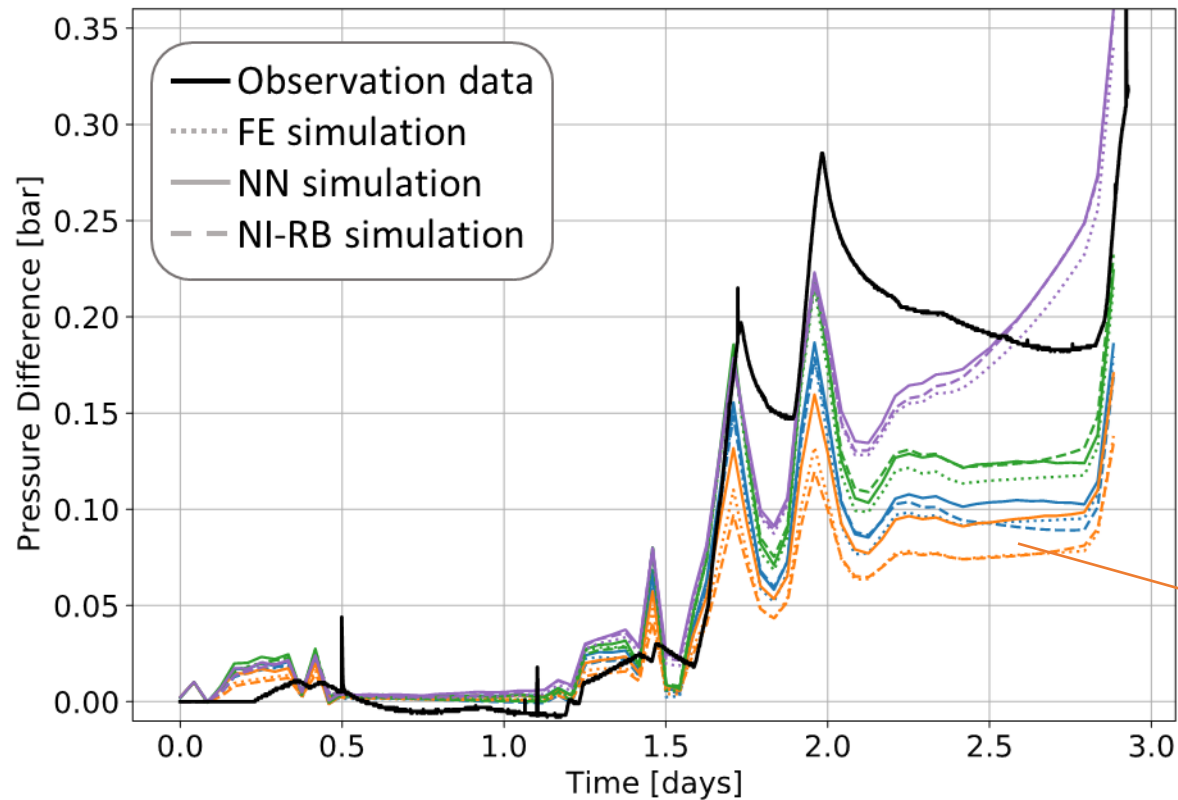
PARAMETER ESTIMATION

Improved “best-fit” model with quantified uncertainties



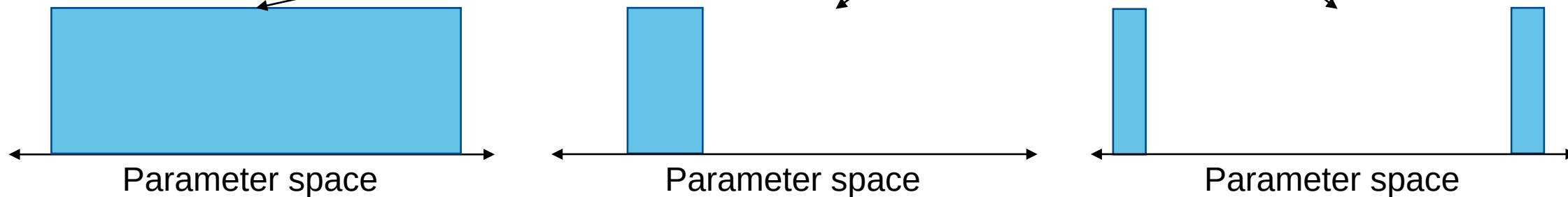
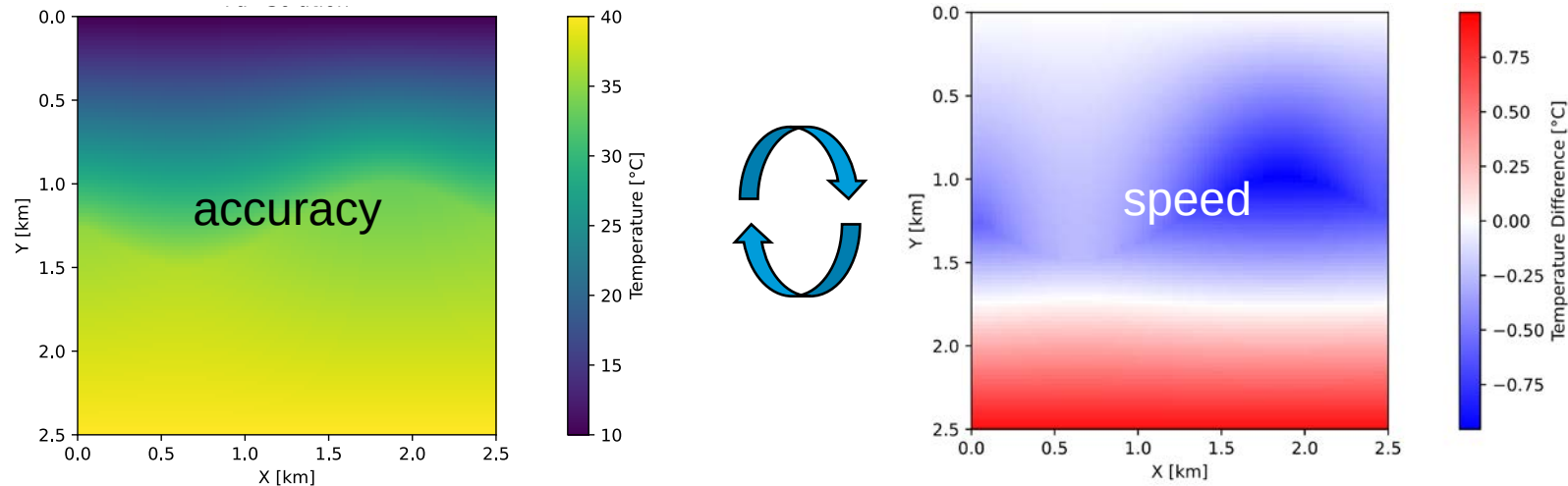
Degen et al. (2022b)

RELIABLE AND INTERPRETABLE MODELS



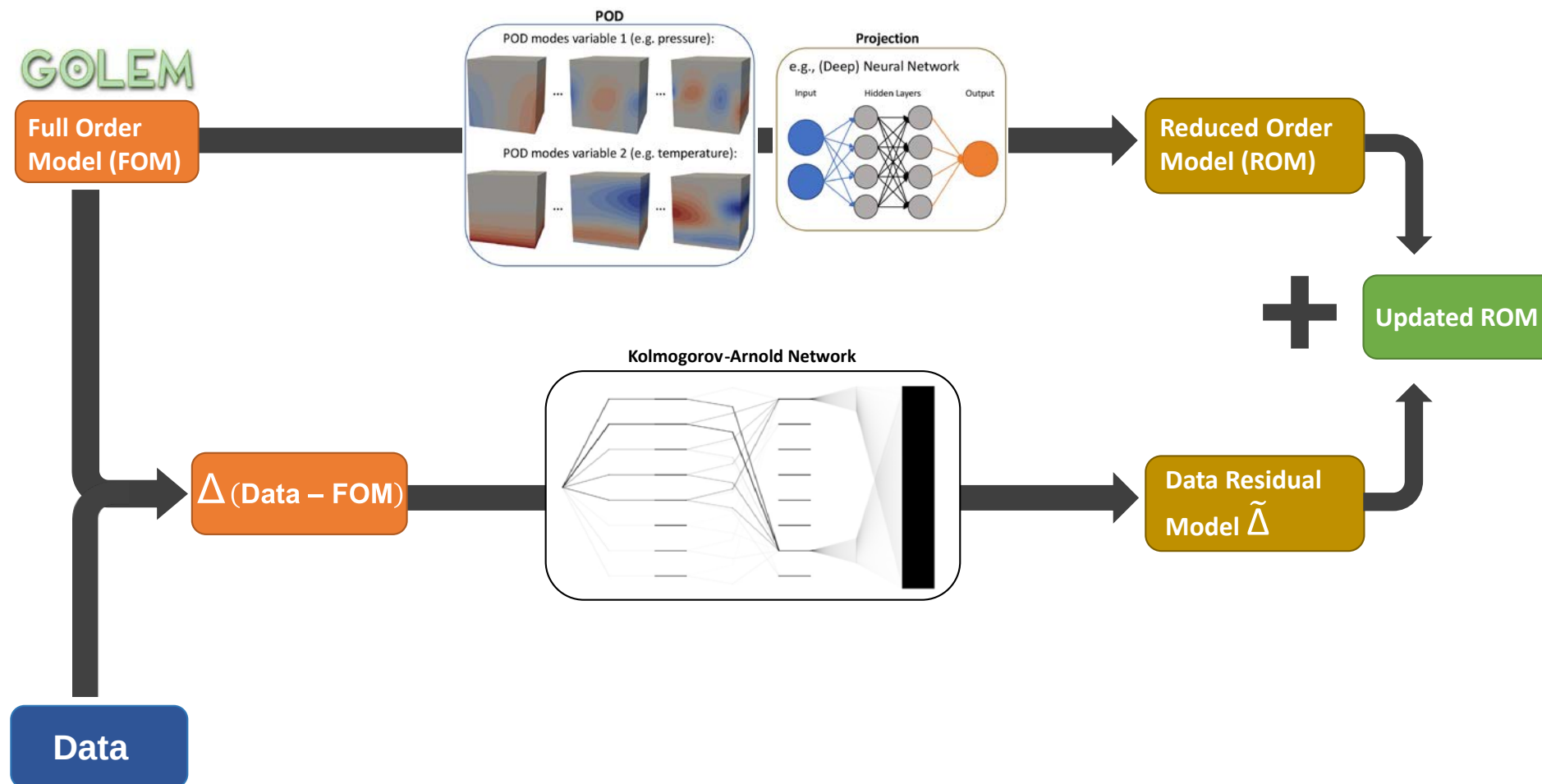
Degen et al. (2024)

HIERARCHICAL APPROACHES

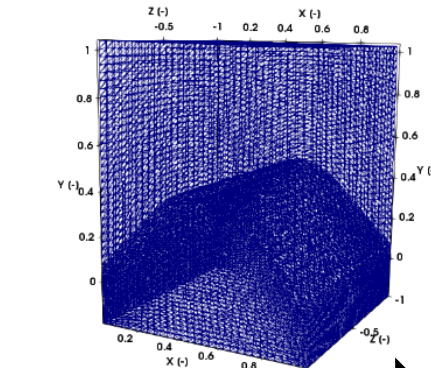
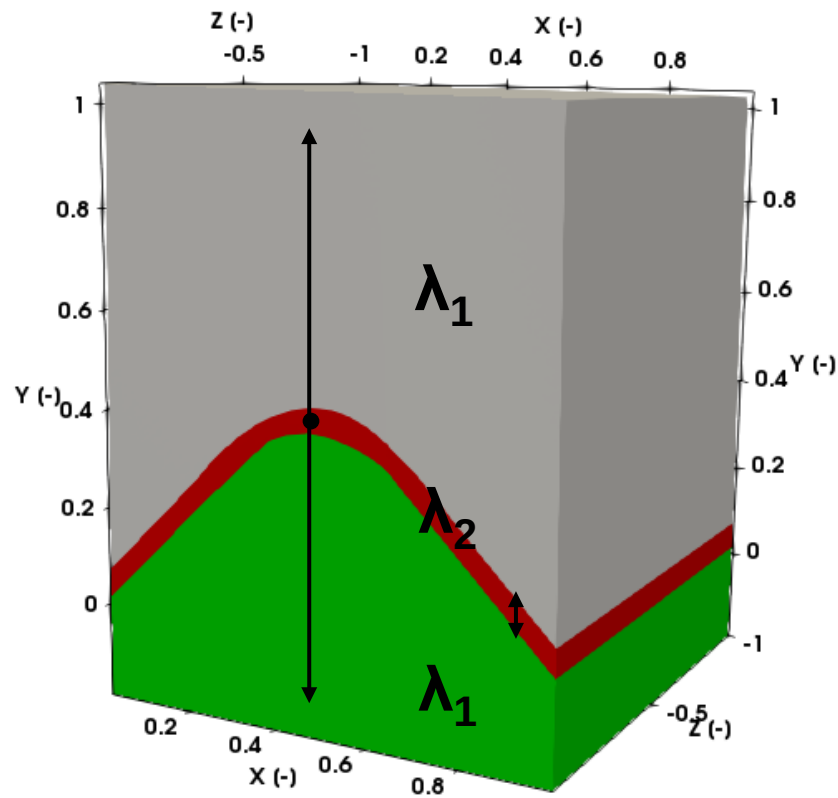


Gruzdeva et al. (2026, in prep.)

DATA INCOPERATION

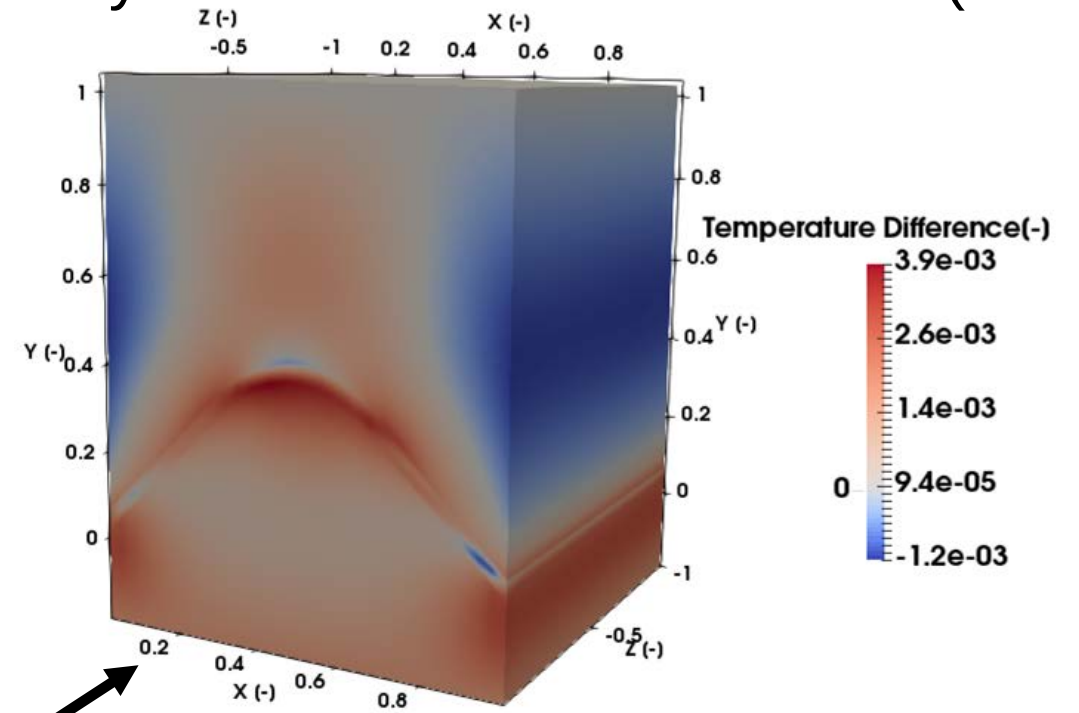


FLEXIBILITY



CNN

PySubdiv → Moulaeifard et al. (2022)

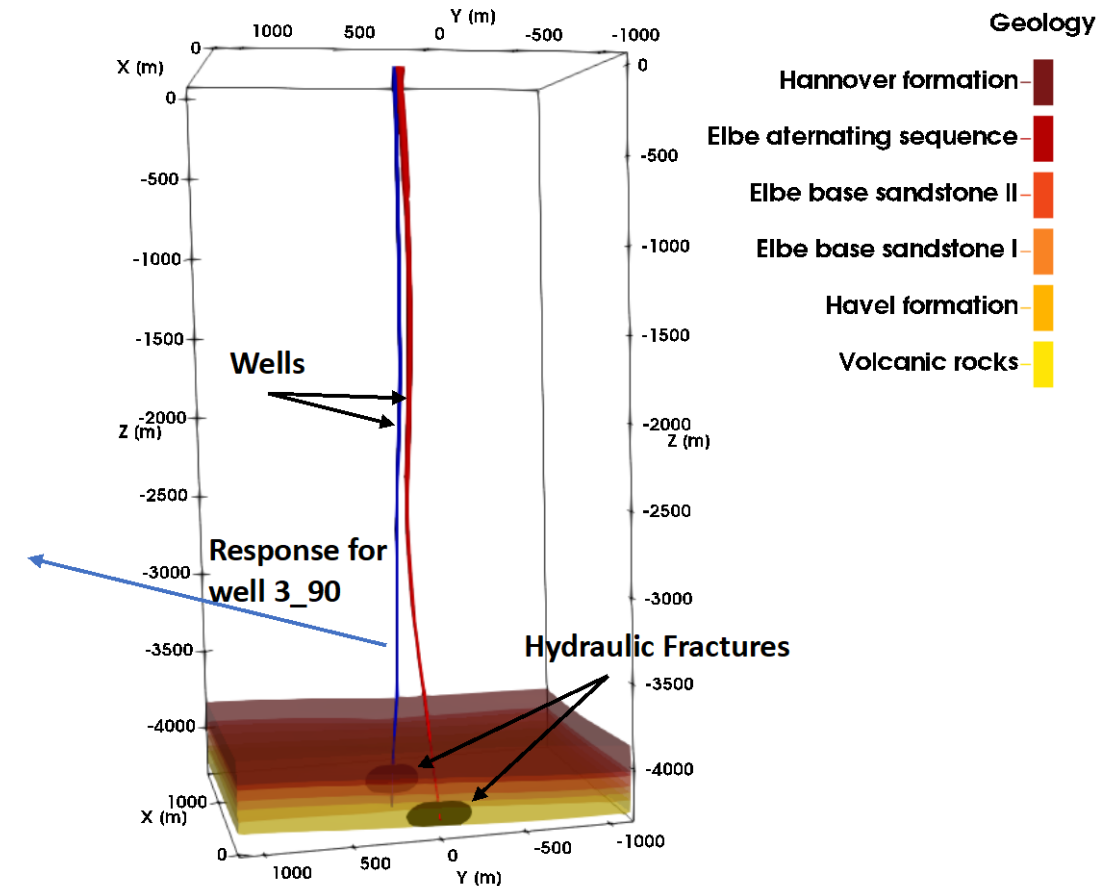


NI-RB

Degen et al. (2026, in prep.)

CONCLUSIONS

- The non-intrusive reduced basis method allows an efficient construction of reliable surrogate models for a wide range of multi-physics applications
- Enabling a digital twin integration and extensive parameter estimation studies
- Resulting in interpretable models
- Computational Gain up to 10^7



Degen et al. (2022b)

REFERENCES

- 1) Degen, D., Veroy, K., Cacace, M., Scheck-Wenderoth, M., & Wellmann, F. (2020a, March). How well do we know our models?. In EGU General Assembly Conference Abstracts (p. 10482).
- 2) Degen, D., Veroy, K., & Wellmann, F. (2020b). Certified reduced basis method in geosciences. *Computational Geosciences*, 24(1), 241-259.
- 3) Degen, D., Veroy, K & Wellmann, F. (2022a). Uncertainty quantification for basin-scale conduction models. *Scientific Reports*, 12(1),1-10.
- 4) Degen, D., Cacace, M. & Wellmann, F. (2022b). 3D multi-physics uncertainty quantification using physics-based machine learning. *Sci Rep* **12**, 17491, <https://doi.org/10.1038/s41598-022-21739-7>.
- 5) Degen, D., Caviedes Voullième, D., Buiter, S., Hendricks Franssen, H. J., Vereecken, H., González-Nicolás, A., & Wellmann, F. (2023). Perspectives of physics-based machine learning strategies for geoscientific applications governed by partial differential equations. *Geoscientific Model Development*, 16(24), 7375-7409.
- 6) Degen, D., Cacace, M., & Wellmann, F. (2024). Exploring physics-based machine learning for geothermal applications. *49th Work Geotherm Reserv Eng*, 1-9.
- 7) Degen, D., & Wellmann, F. (2026). Guidelines for Conducting Sensitivity Analyses in Process Simulations for Solid Earth Geosciences. *Mathematical geosciences*, 1-45.
- 8) Hesthaven, J. S., & Ubbiali, S. (2018). Non-intrusive reduced order modeling of nonlinear problems using neural networks. *Journal of Computational Physics*, 363, 55-78.
- 9) Moulaeifard, M., Bernard, S., & Wellmann, F. (2022). PySubdiv: open-source geological modelling and reconstruction by non-manifold subdivision surfaces. *EGUsphere*, 2022, 1-21.
- 10) Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation*, 55(1-3), 271-280.
- 11) Wainwright, H. M., Finsterle, S., Jung, Y., Zhou, Q., & Birkholzer, J. T. (2014). Making sense of global sensitivity analyses. *Computers & Geosciences*, 65, 84-94.