SMART-Phase 2

Science-informed Machine Learning to Accelerate Real Time (SMART) Decisions in Subsurface Applications

Alexandre Tartakovsky

University of Illinois Urbana-Champaign

Physics-Informed Machine Learning

U.S. DEPARTMENT OF ENERGY

NATIONAL ENERGY TECHNOLOGY LABORATORY

Disclaimer

This project was funded by the U.S. Department of Energy, National Energy Technology Laboratory, in part, through a site support contract. Neither the United States Government nor any agency thereof, nor any of their employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

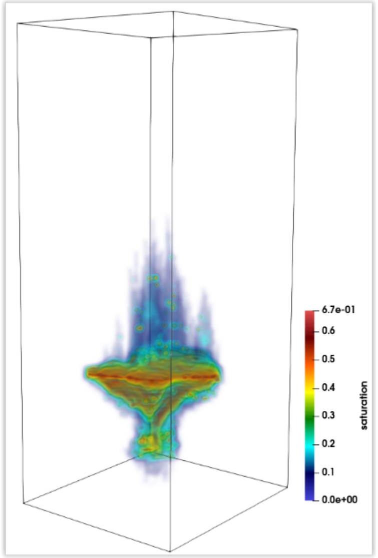
Authors and Contact Information

Alexandre Tartakovsky, Christian Munoz Oro, and Wenhao Chen

University of Illinois Urbana Champaign, Urbana, IL 61802, USA

Why Physics-Informed AI/ML?

Do we need physics-informed AI/ML to predict CO2 plumes (create CO2 plume images) under different conditions (reservoir properties, injection rates, etc)?



Eclipse simulation of CO2 plume at the IBDP site

What Copilot (ChatGPT) can do?

What Copilot (ChatGPT) can do?

Request: draw a mammoth

What Copilot (ChatGPT) can do?

Request: draw a mammoth

Request: draw a real mammoth

Which images are more anatomically correct?

What Copilot (ChatGPT) can do?

Request: draw a CO2 plume

Request: draw a real CO2 plume

What Copilot (ChatGPT) can do?

Request: draw a CO2 plume

Request: draw a real CO2 plume

Request: draw a subsurface co2 plume resulting from co2 sequestration

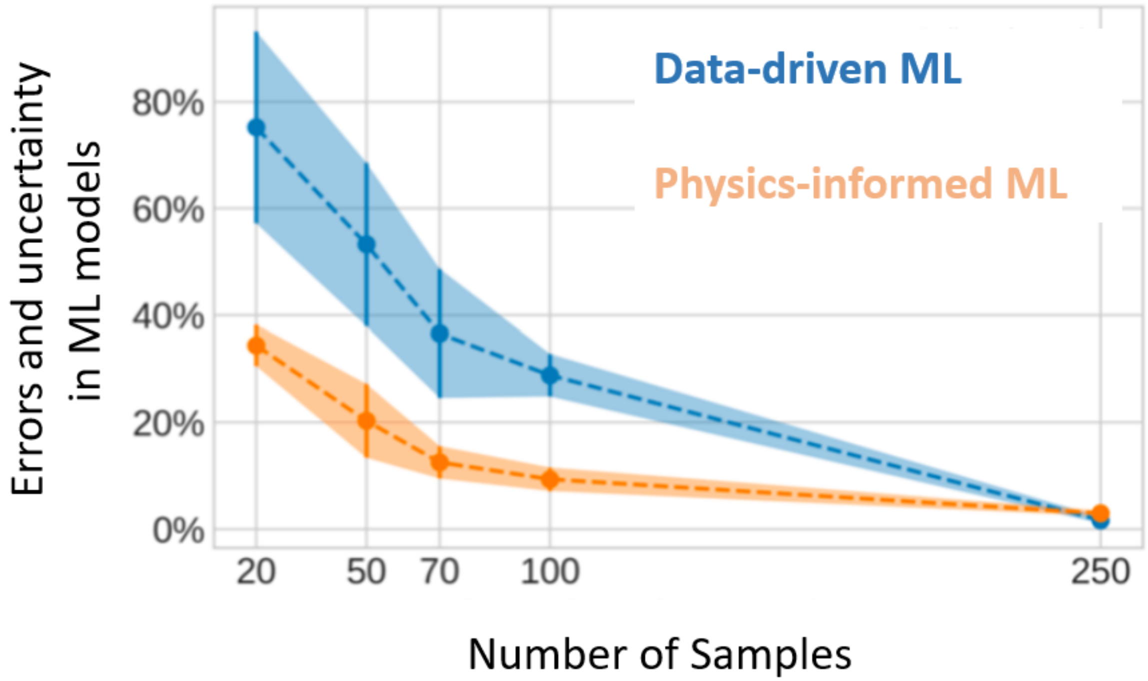
Copilot

Creating an image of a subsurface CO2 plume resulting from CO2 sequestration involves understanding the complex processes that occur underground. When CO2 is injected into…

While I can’t draw or provide images directly, I can describe the expected appearance based on scientific research and simulations. ”The real challenge lies in accurately capturing ….”

Why Physics-Informed ML?

- Enable machine learning with small datasets
- Reduce Uncertainty
- Guide ML methods and choose optimal architecture



Tartakovsky et al, Physics-Informed Deep Neural Networks for Learning Parameters and Constitutive Relationships in Subsurface Flow Problems, WRR, 2020.

Illinois Basin Decatur Project (IBDP) Case Study

- Dataset size: 100 simulations computed on 1.7M node mesh and 50 timesteps

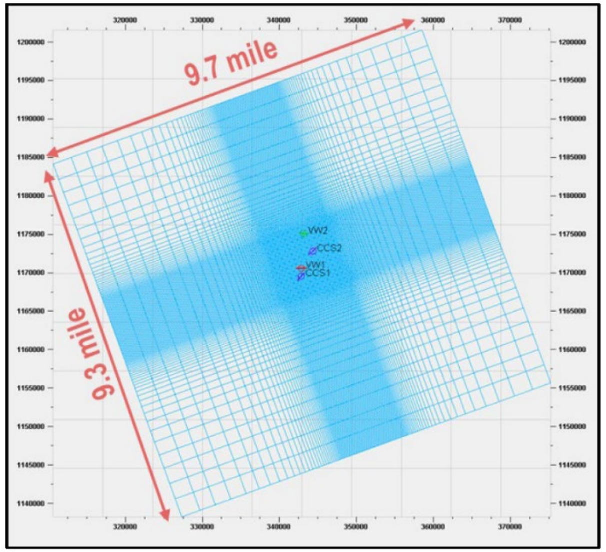
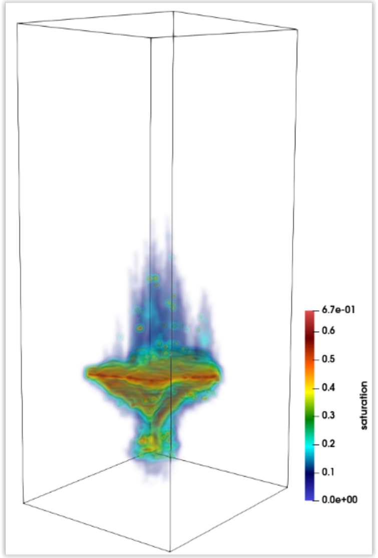


Figure 67. Dynamic model domain and tartan grid.

X-Y Mesh



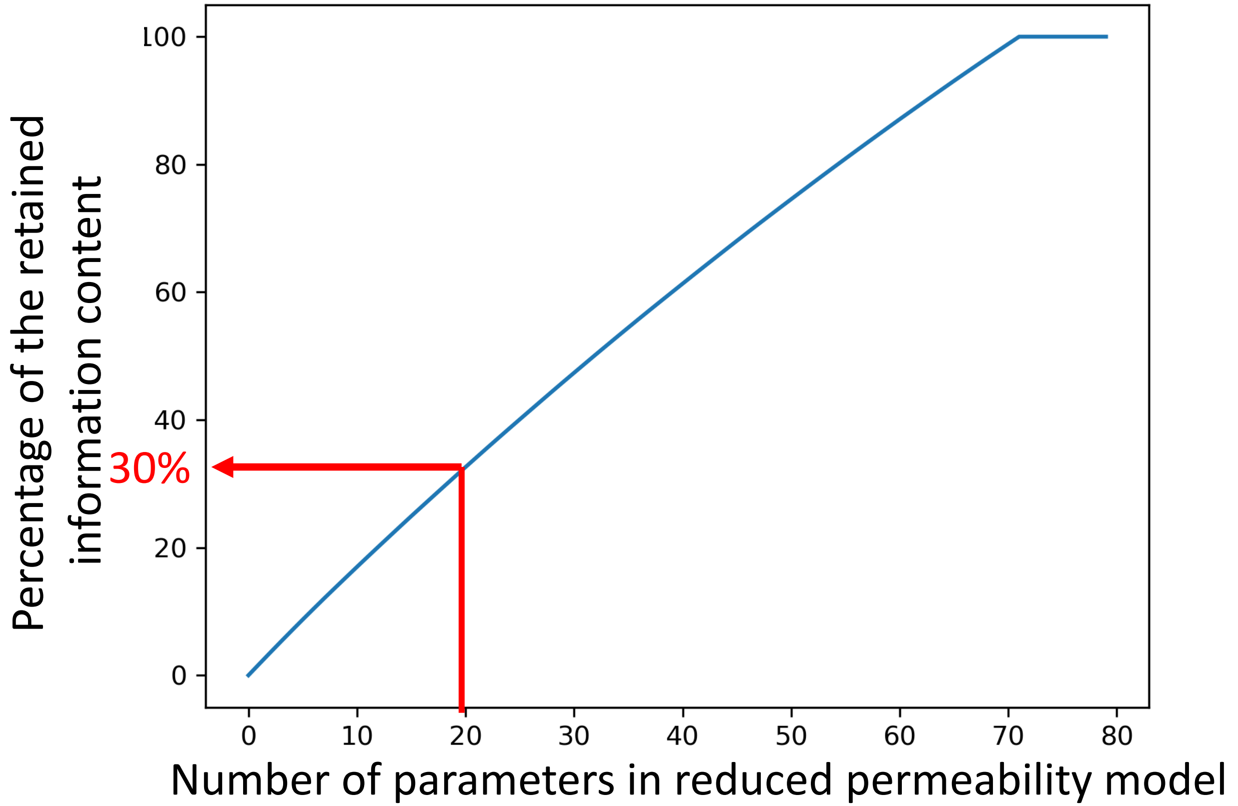
CO2 saturation

- Develop a surrogate model that takes 1.7M permeability values as inputs and produces 1.7M x 50 pressure and saturation values as outputs

Reducing Dataset Dimensionality is Essential but Might Not Suffice

- The regression rule-of-thumb: the number of parameters in the regression model should be 5 to 10 times smaller than the number of samples.
  - With 100 samples, the number of ML model parameters should not exceed 20
  - The simplest (linear) regression model with 20 model parameters has 19 inputs. (A deep neural network model with 19 inputs might have1000 parameters)
- Are 19 inputs enough to describe the IBDP permeability, pressure, and saturation fields?

Linear Dimension Reduction Using Principal Component Analysis



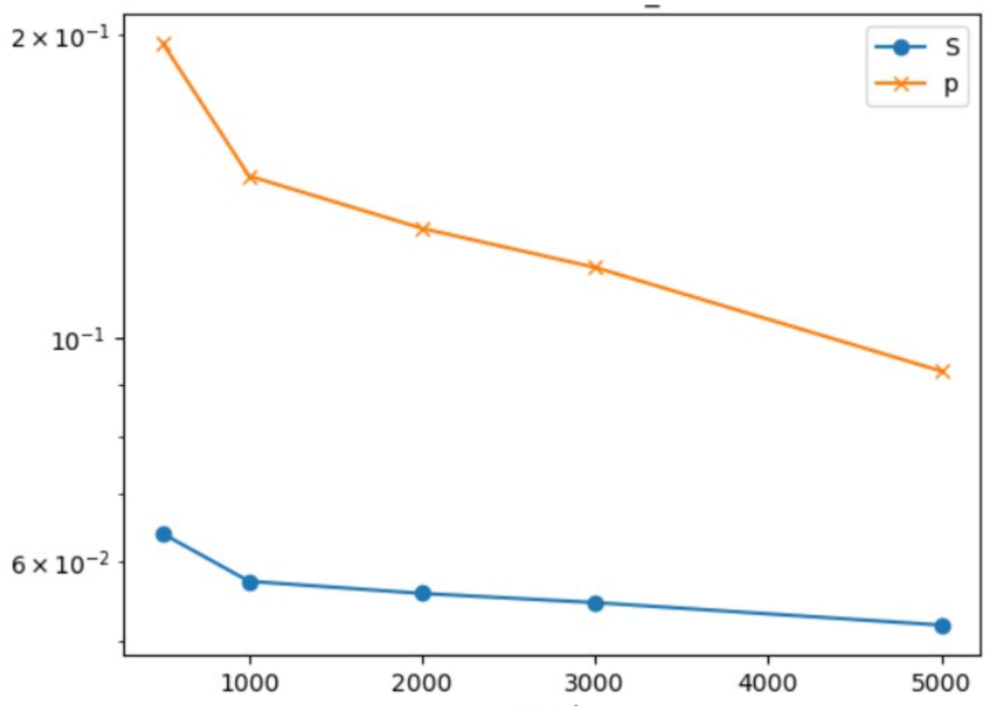
- With 19 reduced dimensions, PCA captures 30% of the total information about the IBDP site
- Linear regression model cannot capture (non-linear) response in the CO2 pressure and saturation (governing physics is highly non-linear).

How to train large regression models with small datasets?

NATIONAL ENERGY TECHNOLOGY LABORATORY

- We can train larger (non-linear) regression models (with more parameters) using the same number of samples, but we must use regularization.
- Standard regularization methods (L2, Tikhonov, or ridge regularization) might not yield informative predictions.
- Physics-based regularization requires the ML model predictions to satisfy governing equations in the least-square sense.
- In the Bayesian framework, regularization (prior assumptions) introduces uncertainty in the predictions.
- Uncertainty quantification is a challenge because of the "curse of dimensionality" in (standard) UQ methods (including Markov Chain Monte Carlo)

Surrogate model with physics constraints



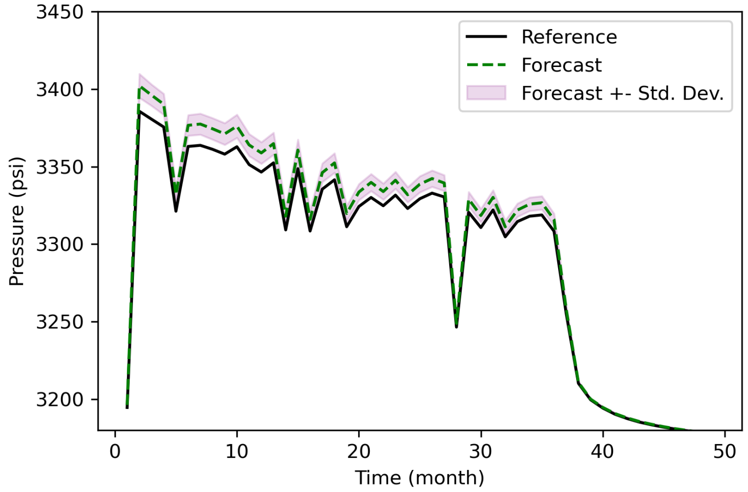
Relative l2 errors in the surrogate model predictions

Number of points where physics is enforced

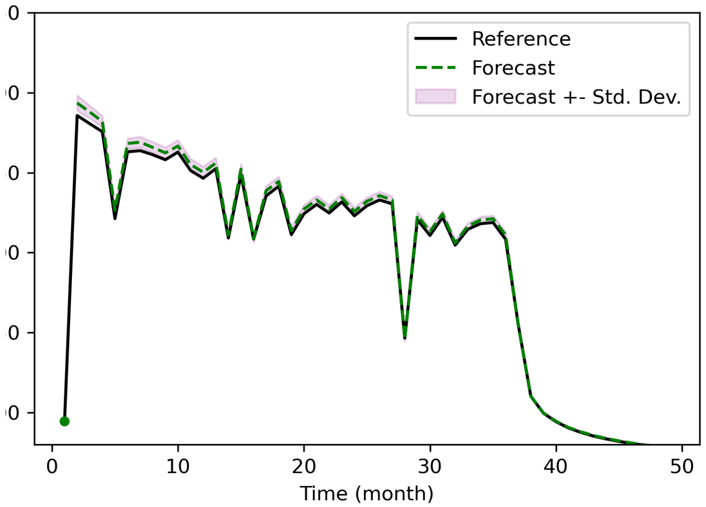
One-dimensional CO2 sequestration test problem. DNN surrogate model with several thousand parameters trained with 100 samples. Physics constraints can reduce the estimation errors by a factor of 2.

UQ in the surrogate model: data assimilation in the pressure forecast

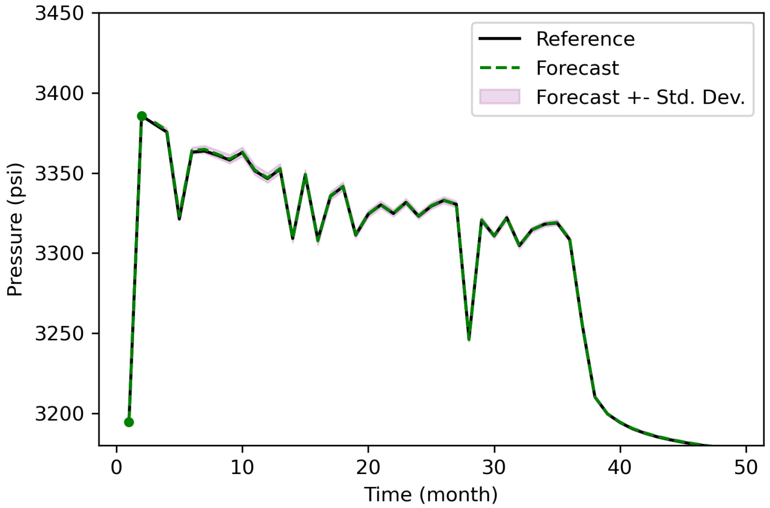
IBDP problem: pressure forecast at an observation well



0 Observations



1 Observation



2 Observations

Surrogate model for pressure forecasting with uncertainty bounds

Questions?

NETL Resources

VISIT US AT: **www.NETL.DOE.gov**

@NETL_DOE

@NETL_DOE

@NationalEnergyTechnologyLaboratory