SMART-CS Initiative

Science-informed **M**achine Learning to **A**ccelerate **R**eal **T**ime (SMART) Decisions in Subsurface Applications

Task 5: Field Deployment – Dynamic Storage Reservoir Modeling

FECM Project Review Meeting
August 28 – September 1, 2023

U.S. DEPARTMENT OF ENERGY

Task 5 –Dynamic Storage Reservoir Modeling

Goal: Provide real-time modeling, data assimilation and forecasting to support:

- Field management -- to maximize storage while minimizing pressure buildup
- Induced seismicity risk assessment

GCS Simulation Today:

- Human-labor intensive
- Heuristic exploration
- Slow and non-interactive

- Decision-Driven
- Ensemble Based

SMART Vision:

- Human-labor efficient
- Automated workflows
- Highly interactive

- Decision-Driven
- Ensemble Based *+ ML Acceleration*

Task 5 –Dynamic Storage Reservoir Modeling

Goal: Provide real-time modeling, data assimilation and forecasting to support:

- Field management, to maximize storage while minimizing pressure buildup
- Induced seismicity risk assessment

Today

5.1: Unified simulation platform and data generation ← Part 1

5.2: Rapid physics-based predictive models for flow and geomechanics

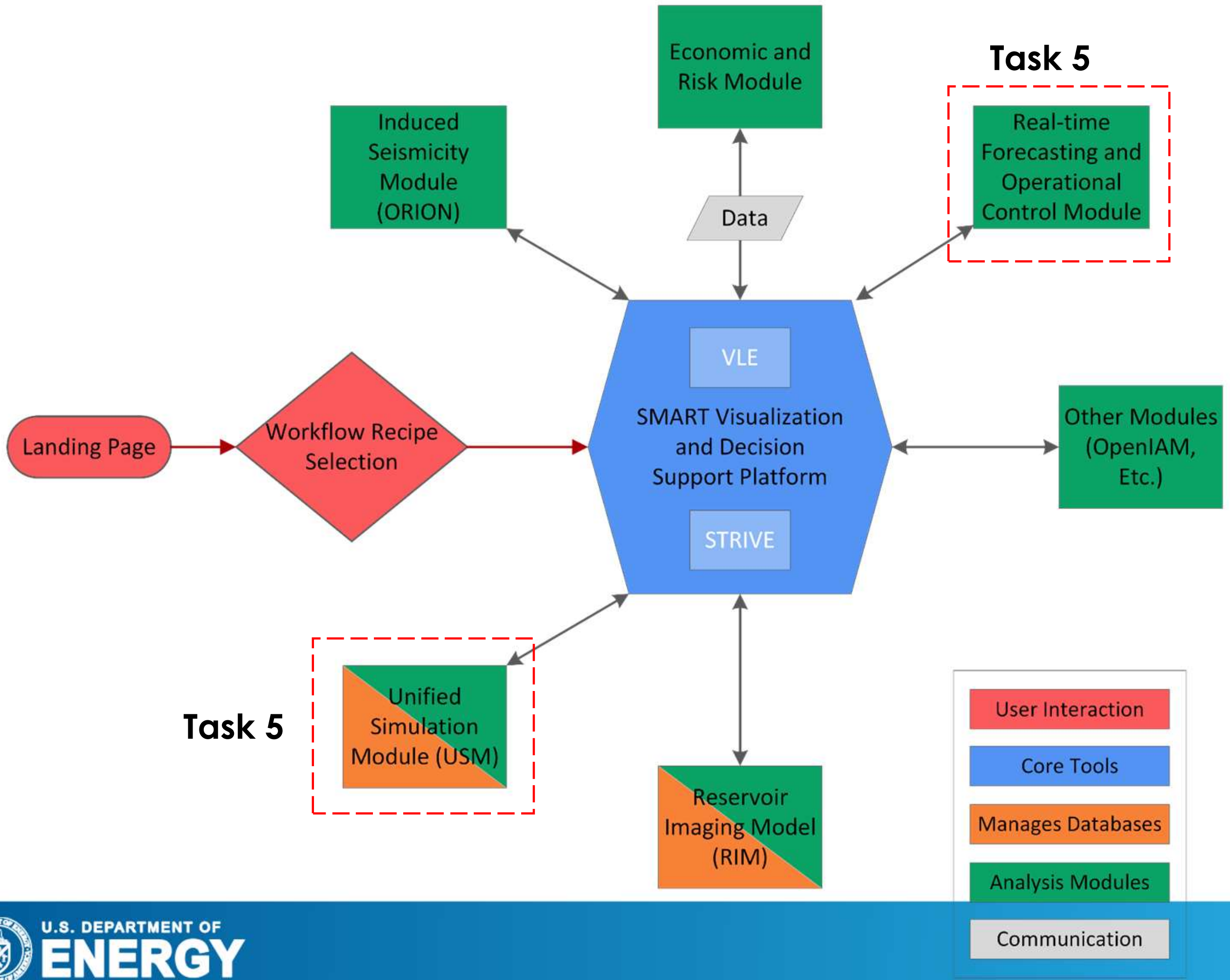
5.3: Machine learning surrogate models ← Part 2

5.4: Rapid data assimilation and history matching ← Part 3

5.5: Optimization of field Parameters

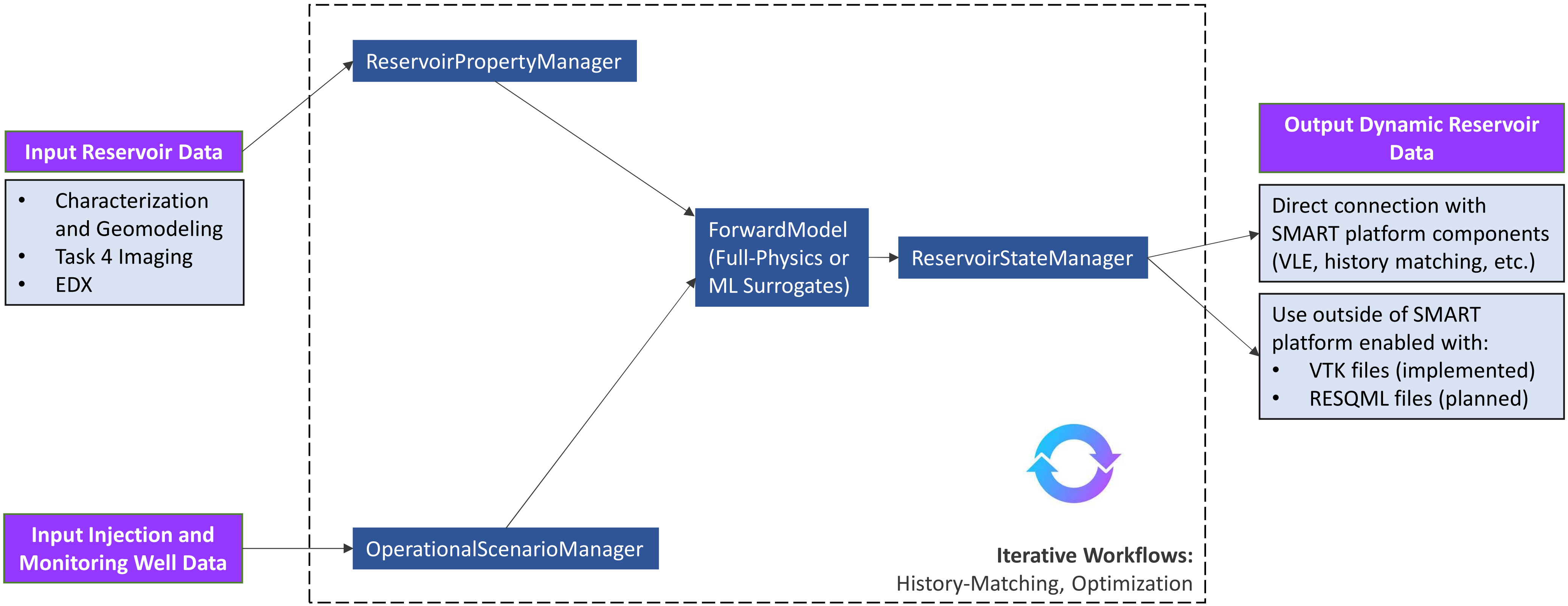
Part 1: Unified Simulation Module

Unified Simulation Module



Objective: Provide a unified way for a user to interact with reservoir simulation data and run simulation workflows

Unified Simulation Module Data Flow



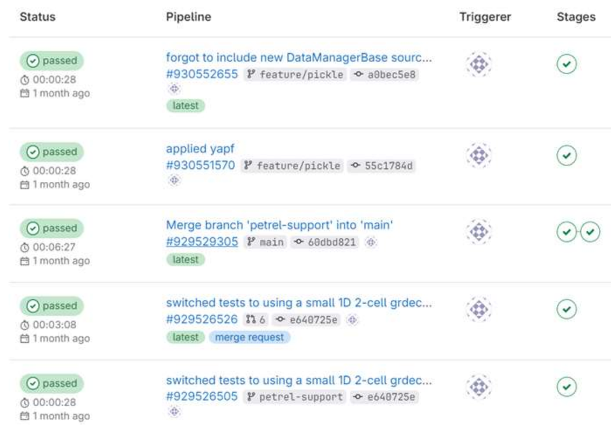
Unified Simulation Module

Goal is a high-quality, shareable, enduring capability

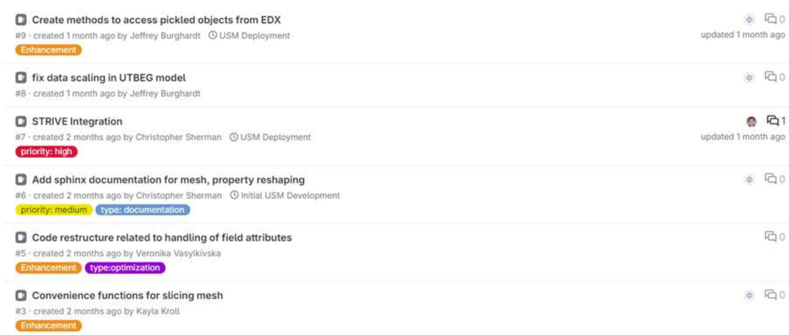
Quality control and documentation

- Code hosted on GitLab
- Installable Python package makes it easy to use
- Automated unit testing suite tests every commit pushed
- Standardized code formatting and style
- Sphinx documentation is automatically built
- Issue and milestone tracking
- *Following recommended software dev practices (Task 3)*

Testing pipeline



Issue tracking

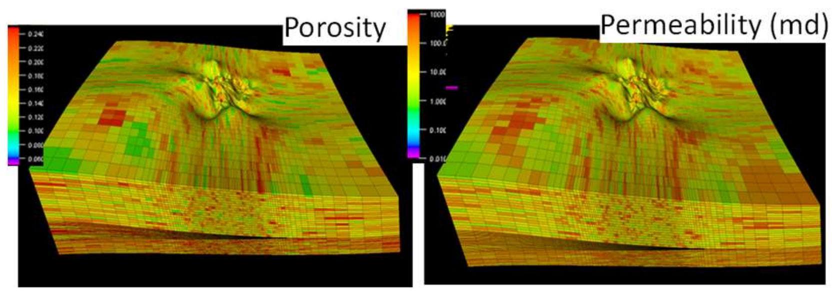


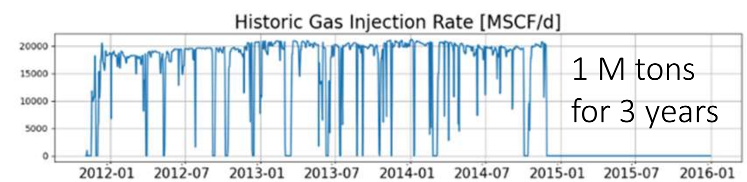
Part 2: ML Surrogate Modeling

ML Input Data

- Monthly pressure and saturation distributions at Illinois Basin Decatur Project (IBDP) site in 100 realizations of permeability and porosity fields (1.73M cells) with actual CO_2 injection rates

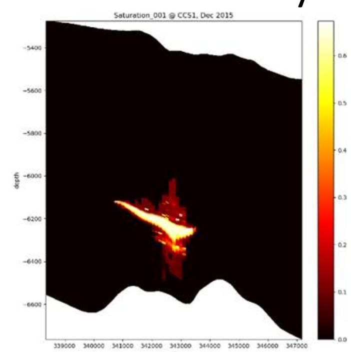
- Training (90 cases) and testing (10 cases)
- Input data
  - Injection rate: (100, 50)
  - Permeability: (100, 126, 125, 110, 3)
  - Porosity: (100, 126, 125, 110)
  - Topology: (100, 126, 125, 110)
- Output data
  - Pressure: (, 50, 126, 125, 110)
  - Saturation: (, 50, 40, 44, 94)
- Well data
  - Injection rates: three perforation zones
  - Monitoring: 6 multi-depth sensors



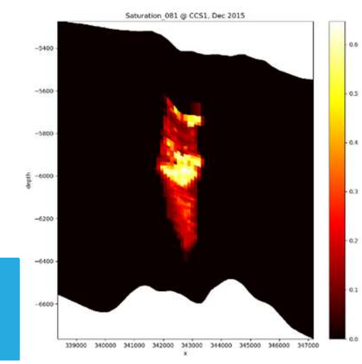


Example of porosity, permeability, and injection rates (input to ML models) & examples of CO_2 saturation distribution at 1 year after the end of injection (Eclipse)

80 cases with open fault horizontally



20 cases with closed fault horizontally



ML Models

Three primary goals:

- Computational efficiency to handle IBDP data (1.73M cells, 50 time steps, 100 realizations)
- Prediction accuracy
- Flexibility associated with input, output, portability, and potentially transfer learning

ORG	ML Method	Pressure RMSE (psi)	Saturation RMSE (-)*	Note
UT-BEG	UNet-MLP	<2	~0.016	Relatively big model (122M parameters, 23.6 hr training on 2 GPUs), handling full IBDP data
ORNL	Autoencoder-MLP (AE-MLP)	~20-25	~0.018	Latent space based approach, 2D slice model for pressure
SNL	Modified DeepONet with subsampling (DeepONet)	~2	~0.018	Subsampling for computational efficiency (~ 1hr training on 1 GPU & 2.2M parameters), handling full IBDP data
LANL	Fourier Neural Operator (FNO-1)	~5	~0.015	2D input due to data size on single GPU
LLNL	Fourier Neural Operator (FNO-2)	~4	~0.015	32 GPUs for ML training with full IBDP data (2 & 1 hrs for P &S)
UIUC	Karhunen-Loeve-Deep Neural Network (KL-DNN)	<2	~0.020	Domain needs to be coarsened in both space and time due to big IBDP data

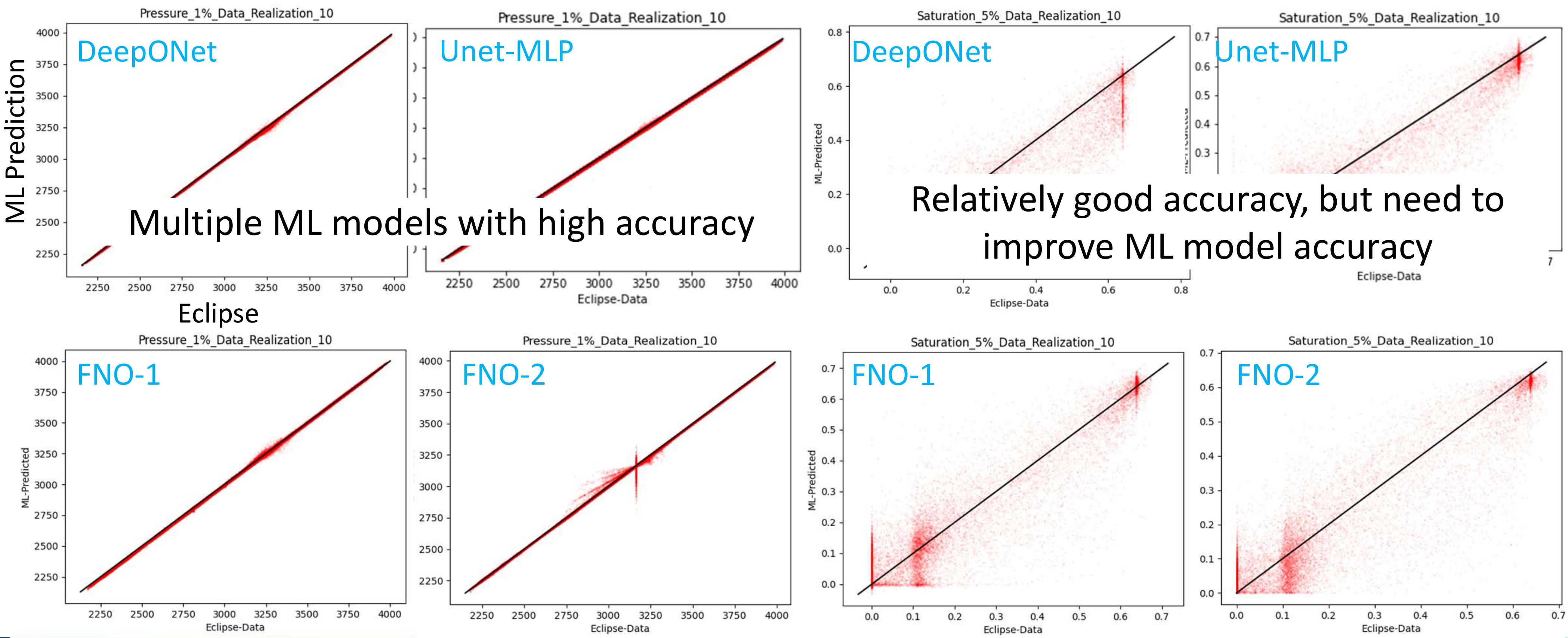
* Saturation evaluation was performed over the central part (~10%) of pressure model domain

* Challenge for saturation ML models: a majority of saturation data are zeros

Pressure & Saturation Prediction (realization 10)

Pressure plot with 1% of entire data

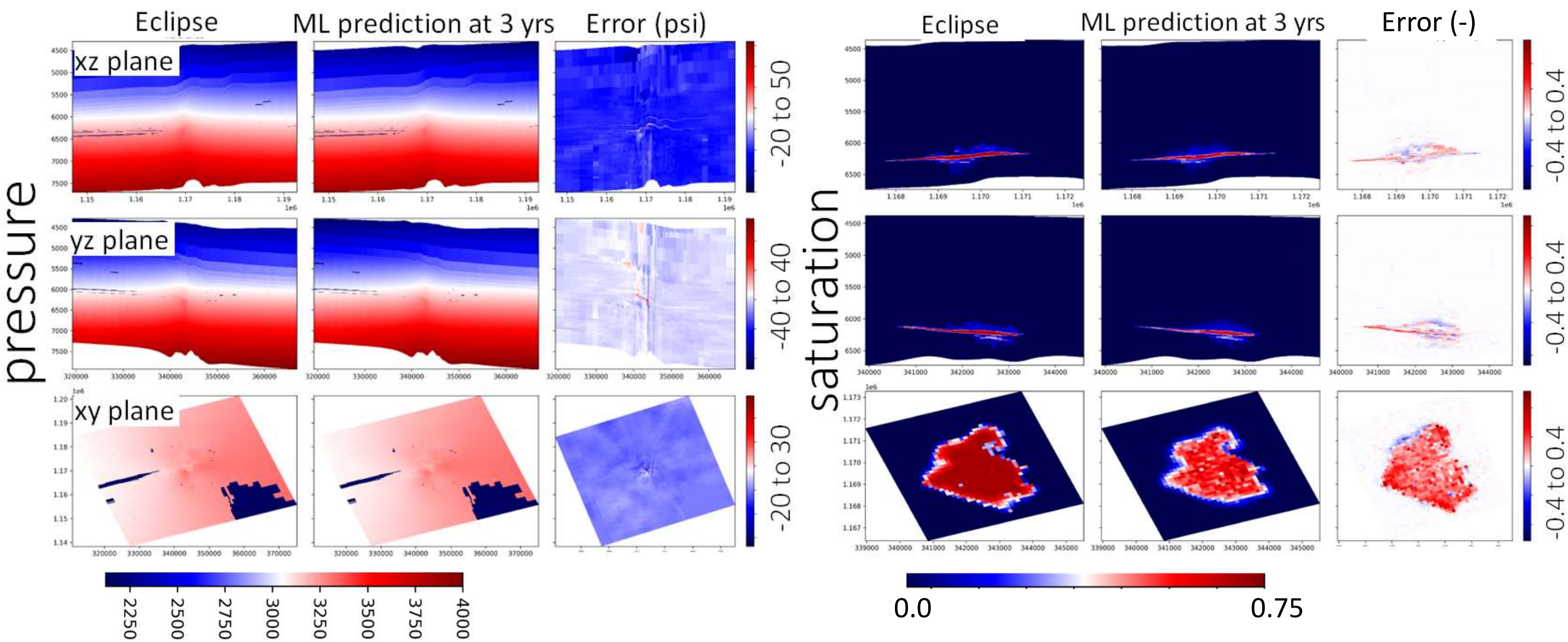
Saturation plot with 5% of entire data



Multiple ML models with high accuracy

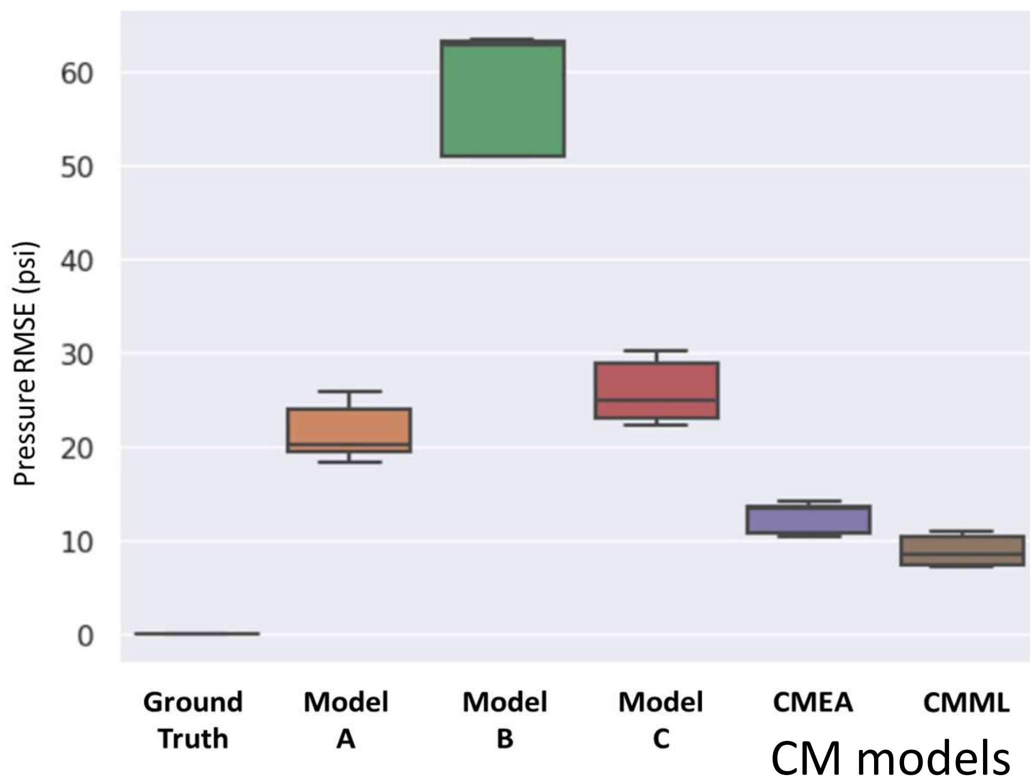
Relatively good accuracy, but need to improve ML model accuracy

Snapshots of Pressure and Saturation



ML Committee Machine (CM)

- **CM**: an ensembled approach to aggregate predictions of multiple ML models into a final decision
- For this demo, a few simple neural network models A-C for pressure at IBDP site are used
- CMs (CMEA and CMML) performed better than individual ML models



Part 3: Accelerated History Matching and Plume Visualization with Machine Learning

Authors and Contact Information

Changqing Yao, Masahiro Nagao, Akhil Datta-Gupta

Petroleum Engineering

Texas A&M University, College Station, TX 77843

Objective and Challenges

- Development and application of ML-assisted tools and workflows for **field-scale application and validation** of geologic carbon storage
  - Rapid forecasting of CO2 plume evolution constrained by observed distributed temperature and pressure data while accounting for data sparsity and geologic uncertainties
- Current Challenges
  - Expensive forward simulation: multiphase, compositional and coupled flow
  - Repeated simulations for model calibration and uncertainty analysis
  - Traditional history matching is time consuming -- often takes weeks/months and is not amenable to real time decision-making

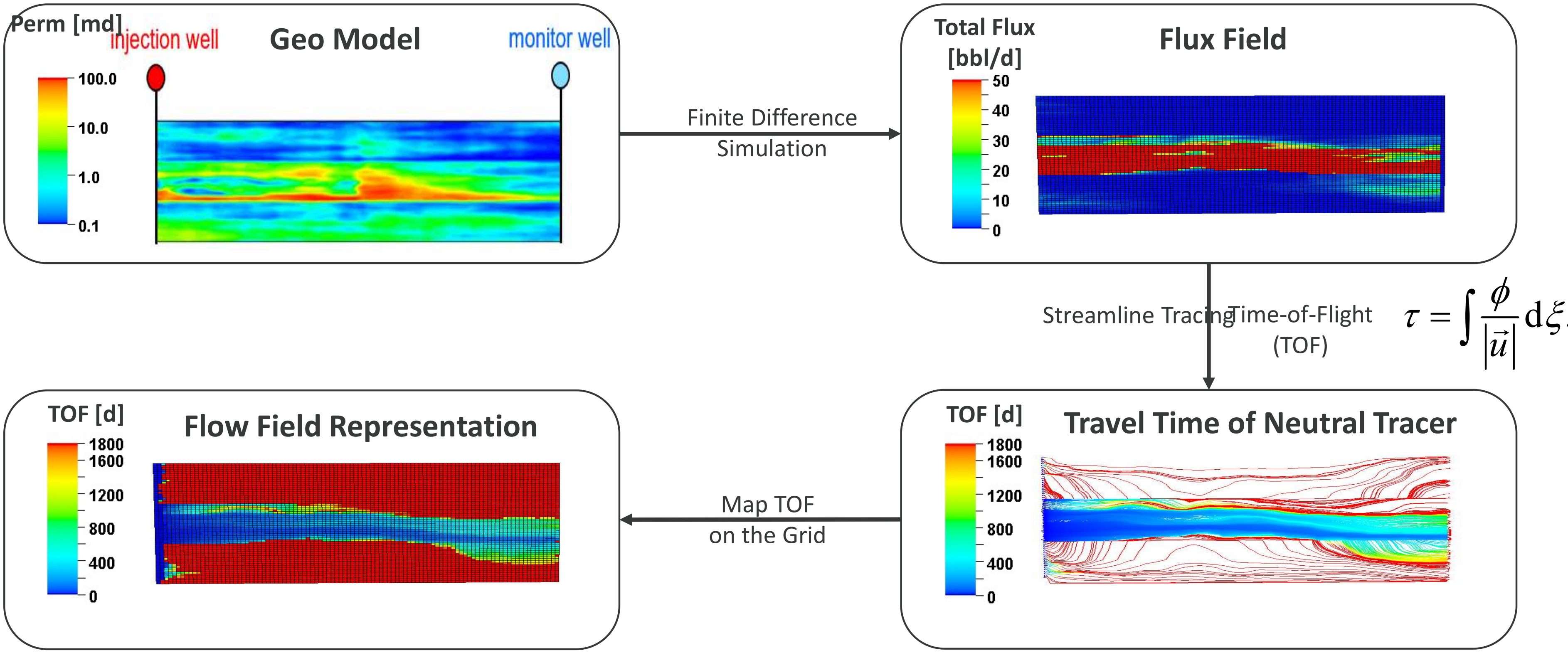
Proposed Workflow: Outline of Steps

- **Dimensionality and computational time reduction for the training data**
  - Single Time-of-Flight map representing CO_2 propagation
- Neural Network Training
  - Variational autoencoder (VAE) to compress time of flight images using latent variables
  - Regression model to estimate autoencoder latent variables based on the monitoring data
- Prediction of CO_2 plume images
  - Estimate Time-of-Flight map from monitoring data (pressure and temperature at the injection and monitoring wells)

Data Dimensionality and Computational Time Reduction:
Single Time-of-Flight Map Representing CO_2 Propagation

NATIONAL ENERGY TECHNOLOGY LABORATORY

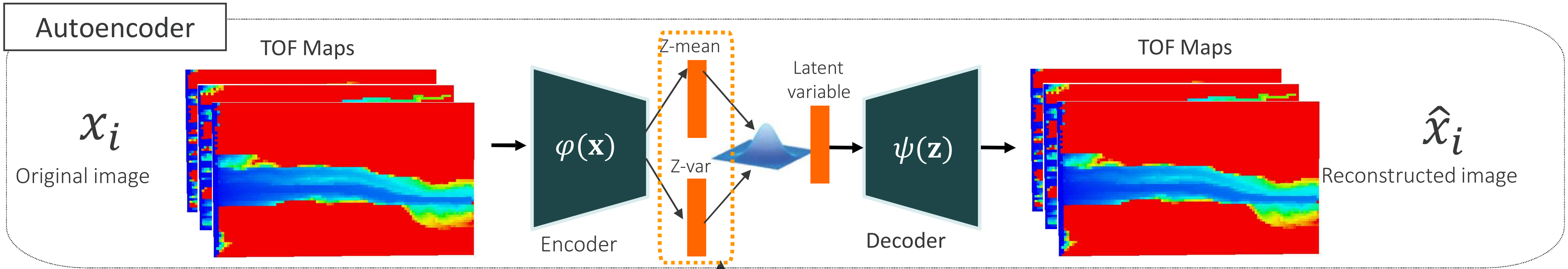
Time of Flight (TOF): Travel time of a neutral tracer along streamlines representing the flow field and fluid transport

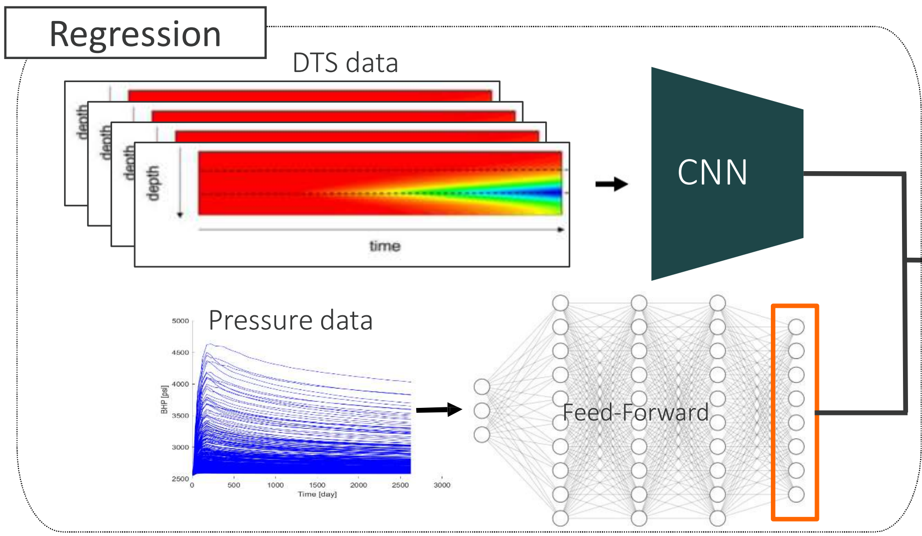


Proposed Workflow: Outline of Steps

- Dimensionality and computational time reduction for the training data
  - Single Time-of-Flight map representing CO2 propagation
- **Neural Network Training**
  - Variational autoencoder (VAE) to compress time of flight images using latent variables
  - Regression model to estimate autoencoder latent variables based on the monitoring data
- Prediction of CO_2 plume images
  - Estimate Time-of-Flight map from monitoring data (pressure and temperature at the injection and monitoring wells)

Neural Network Training





Loss function:

$$L = \underbrace{L_{VAE}}_{\text{Image reconstruction loss}} + \underbrace{\frac{1}{Nz}\sum_{j=1}^{Nz}\left|\sigma_j^2 - \hat{\sigma}_j^2\right|^2 + \frac{1}{Nz}\sum_{j=1}^{Nz}\left|\mu_j - \hat{\mu}_j\right|^2}_{\text{Bottleneck loss}}$$

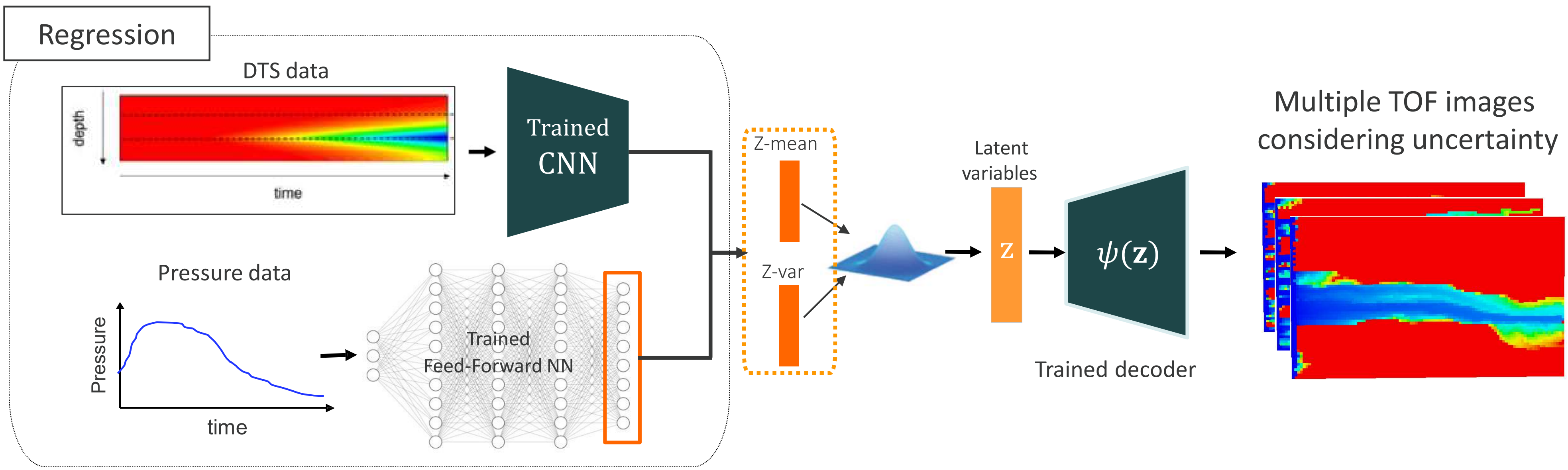
$$L_{VAE} = \frac{1}{N}\sum_{i=1}^{N}|x_i - \hat{x}_i|^2 - \frac{1}{Nz}\sum_{j=1}^{Nz}\frac{1}{2}\left[1 + log\left(\sigma_j^2\right) - \sigma_j^2 - \mu_j^2\right]$$

Proposed Workflow: Outline of Steps

- Dimensionality and computational time reduction for the training data
  - Single Time-of-Flight map representing CO_2 propagation
- Neural Network Training
  - Variational autoencoder (VAE) to compress time of flight images using latent variables
  - Regression model to estimate autoencoder latent variables based on the monitoring data
- **Prediction of CO_2 plume images**
  - Estimate Time-of-Flight map from monitoring data (pressure and temperature at the injection and monitoring wells)

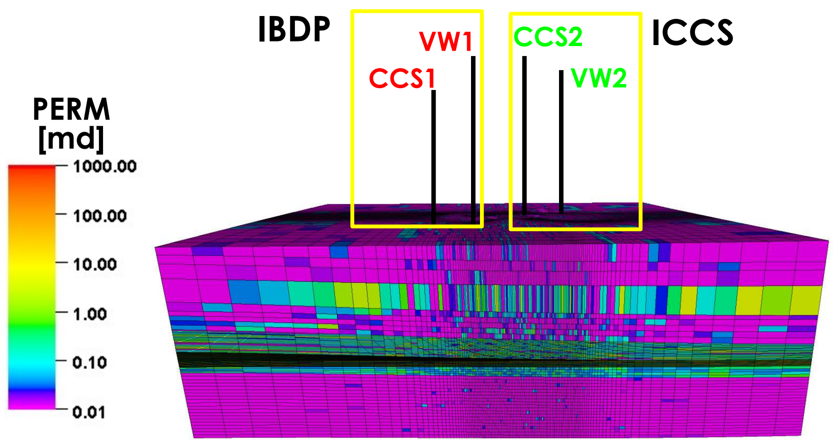
Prediction of CO2 Onset Time Map

Predict the Time-of-Flight map based on the field monitoring measurements

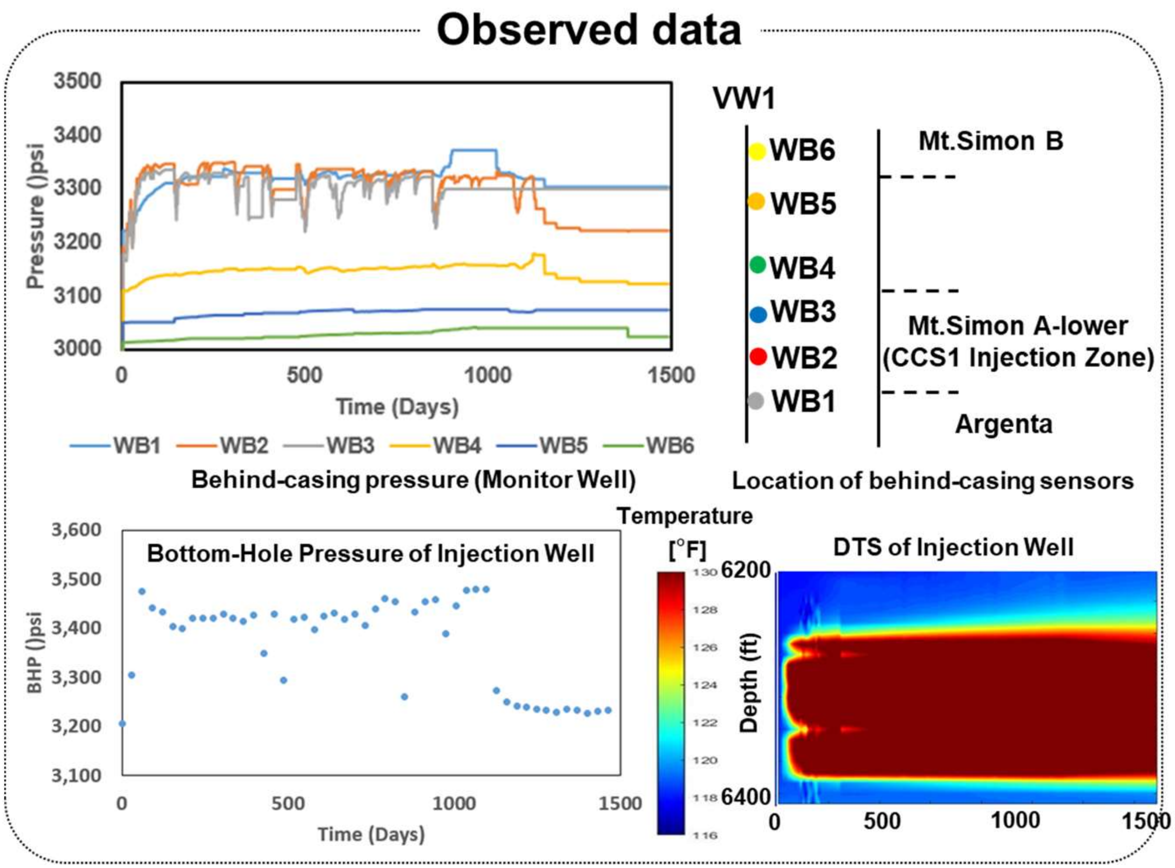


IBDP Model Description and Data Availability

NATIONAL ENERGY TECHNOLOGY LABORATORY

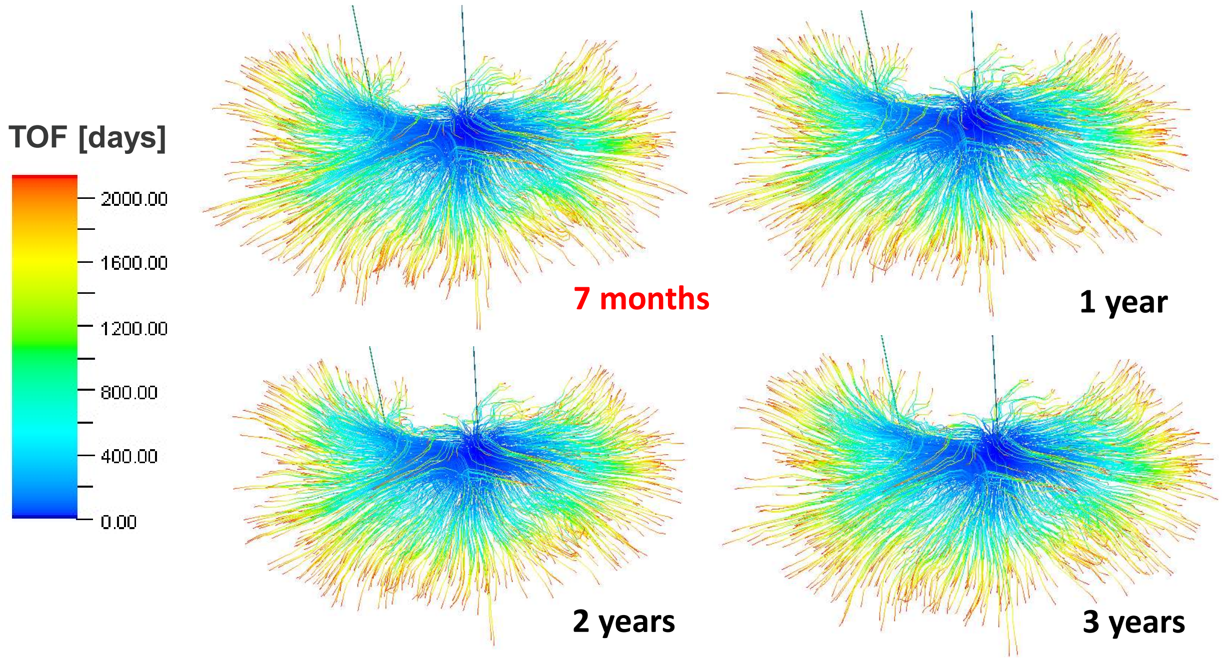


- Grid: 126 * 125 * 110 (1.73 Million Cells)
- ECLIPSE Compositional Model (E300)
- Thermal Option
- CO2STORE Module
- Simulation Period: 2011-2015
- Run Time: 12 hours with 32 Cores Parallel run



Time of Flight Calculations :Validation of Flux Field Stability

- **Challenge**: Run time for a full simulation is too long (12 hrs) and unfeasible for training data generation purposes (hundreds of simulations)
- **Solution**: Calculate time of flight when the flow field is stabilized

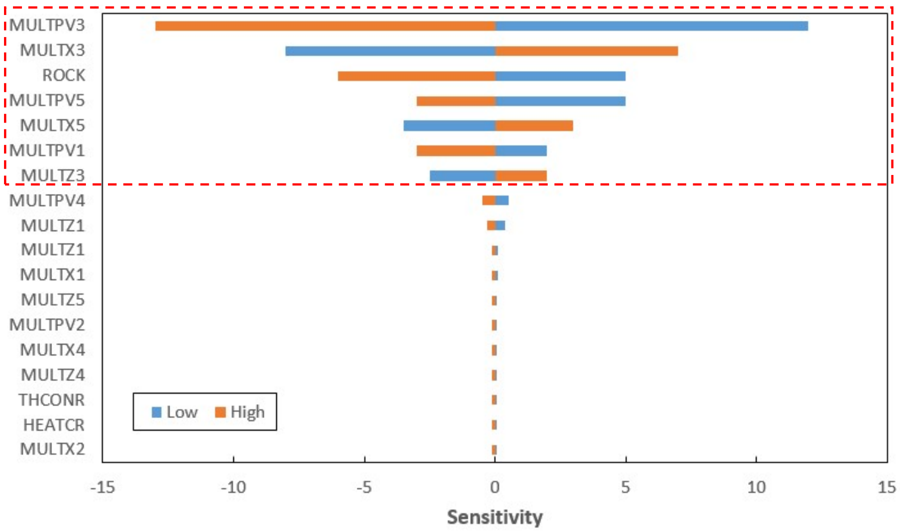


Selected simulation period:
7 months (flux field has been stabilized)

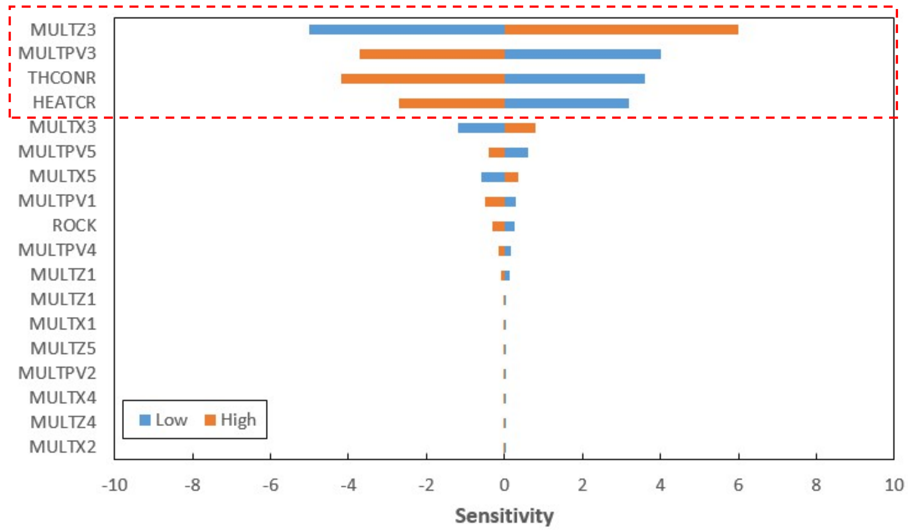
Run Time: 12 hrs → 1 hr

Sensitivity Analysis for Training Data Generation

Pressure Responses (injection well + monitoring well)



DTS (injection well)

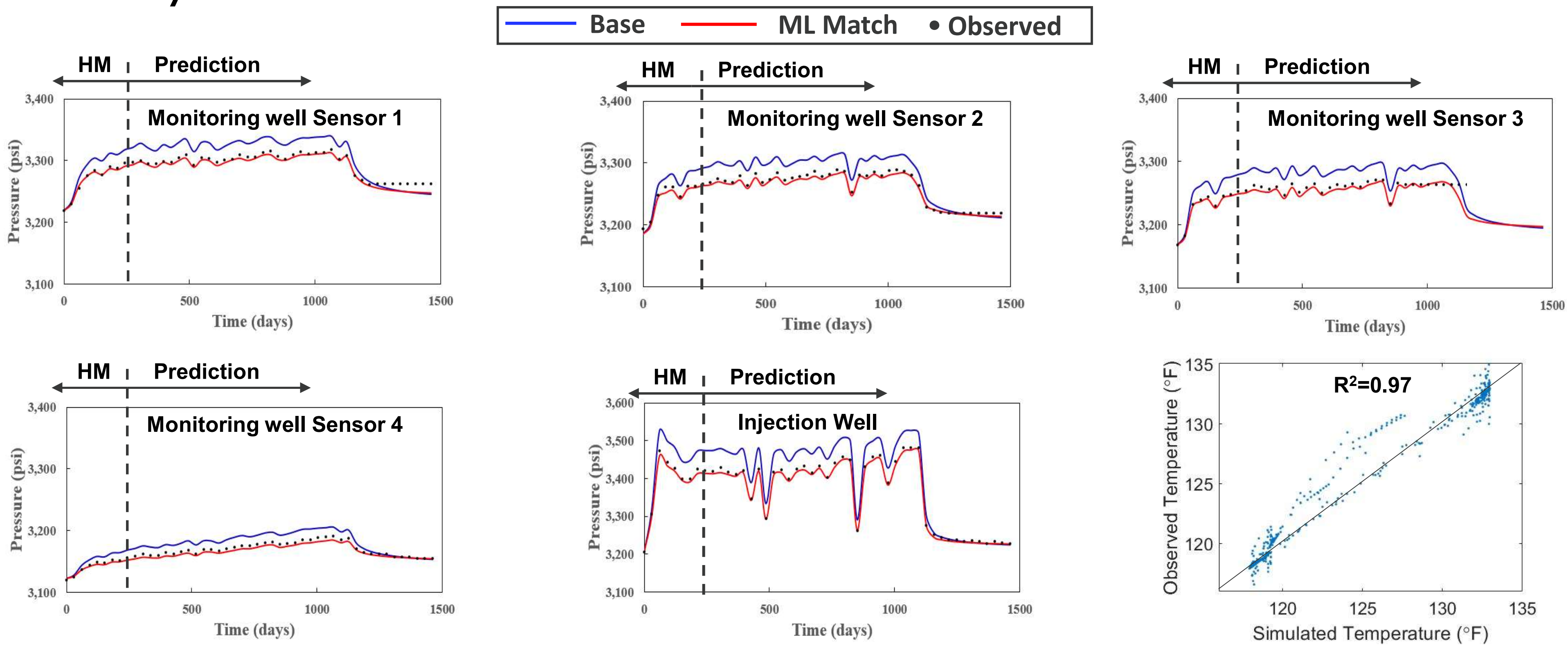


Selected Parameters: MLTPV1/3/5, MULTX3/5, MULTZ3, Rock Compressibility, Rock/Fluid Thermal Conductivity, Rock Heat Capacity

Pressure and Temperature Matching Comparison

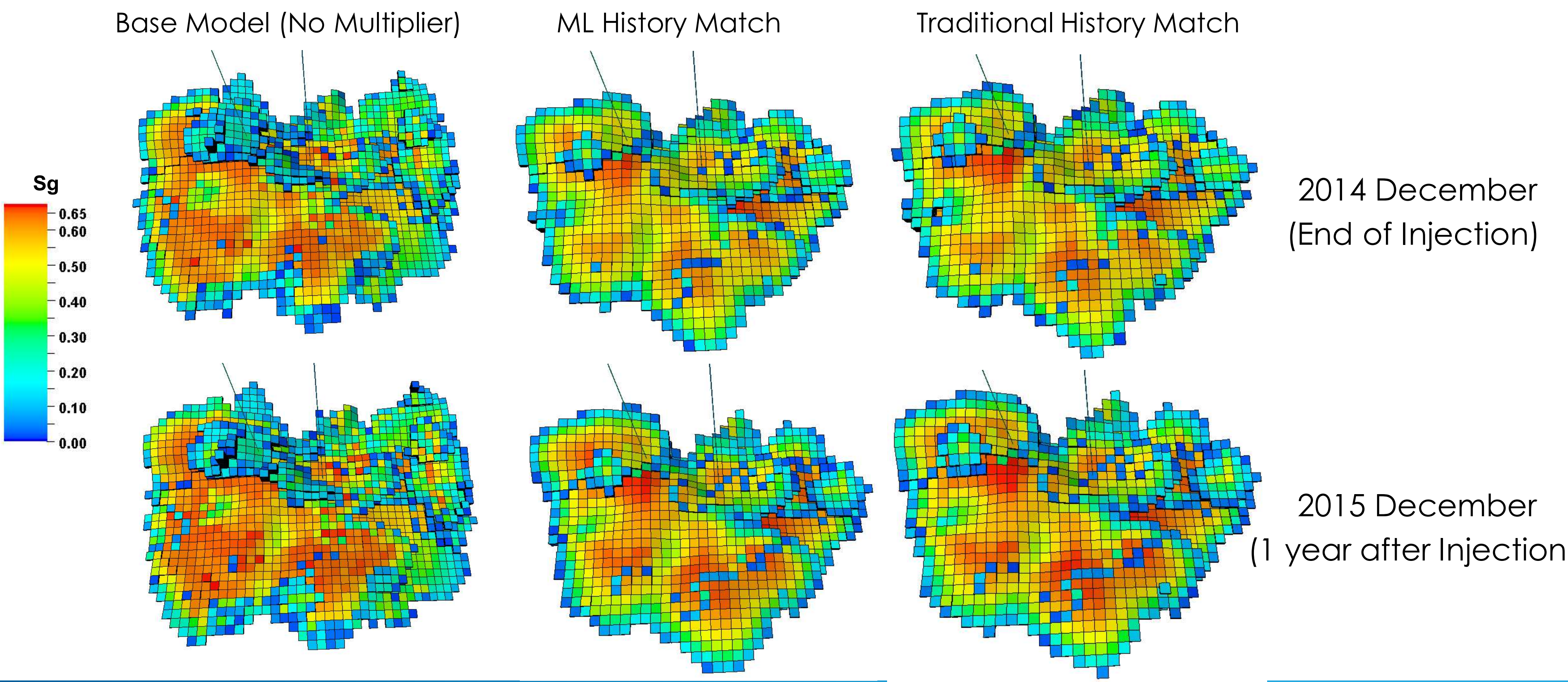
NATIONAL ENERGY TECHNOLOGY LABORATORY

Significant speed-up: 5 hours for training, seconds for model calibration/predictions as opposed to traditional history matching that can take days or weeks.



3D CO_2 Saturation Contour Comparison

Contour generated from CO_2 saturation map with a threshold of 1%



Summary

- Proposed ML assisted workflow and application to the IBDP site shows promising results with orders of magnitude speed up
- Incorporated thermal effects to integrate DTS data and utilized Time-of-Flight to reduce computational time substantially
- Future Opportunities
  - Leverage Oil Industry Experience: Fast Marching Method for Coupled Flow, Streamlines for visualization, Storage/CO2 Sweep Optimization via Rate control

Questions?

Acknowledgements

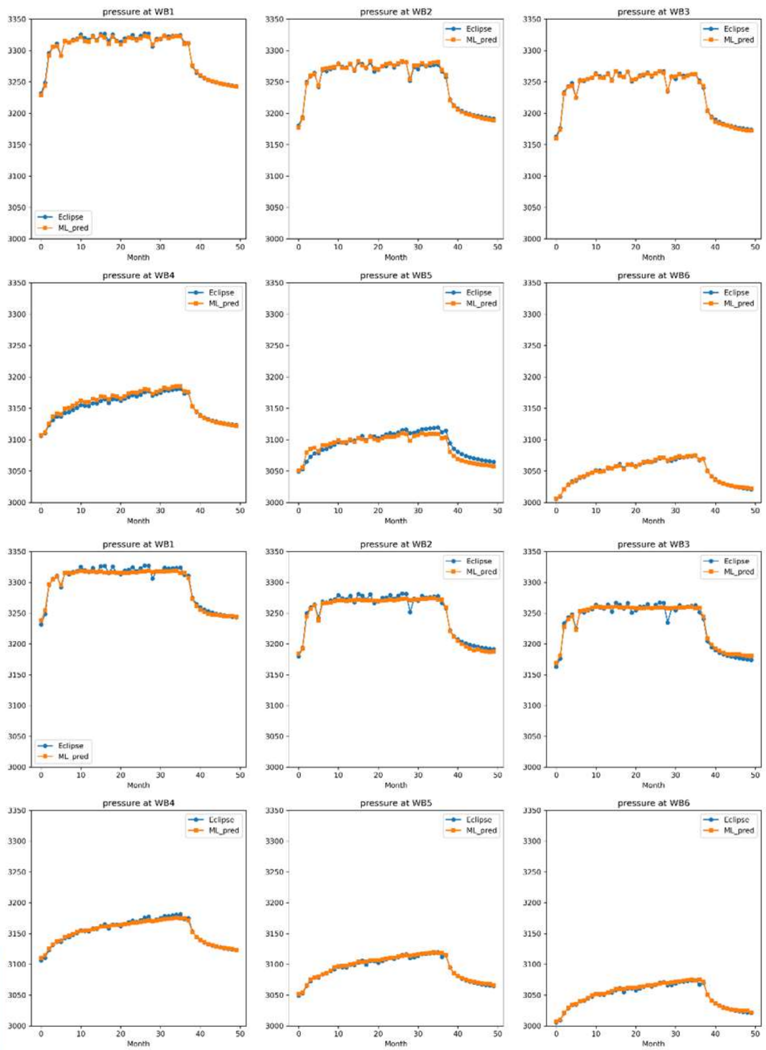
Funding:

- This work was funded by the SMART Initiative through the DOE Office of Fossil Energy and Carbon Management and the Bipartisan Infrastructure Law.

Auspices:

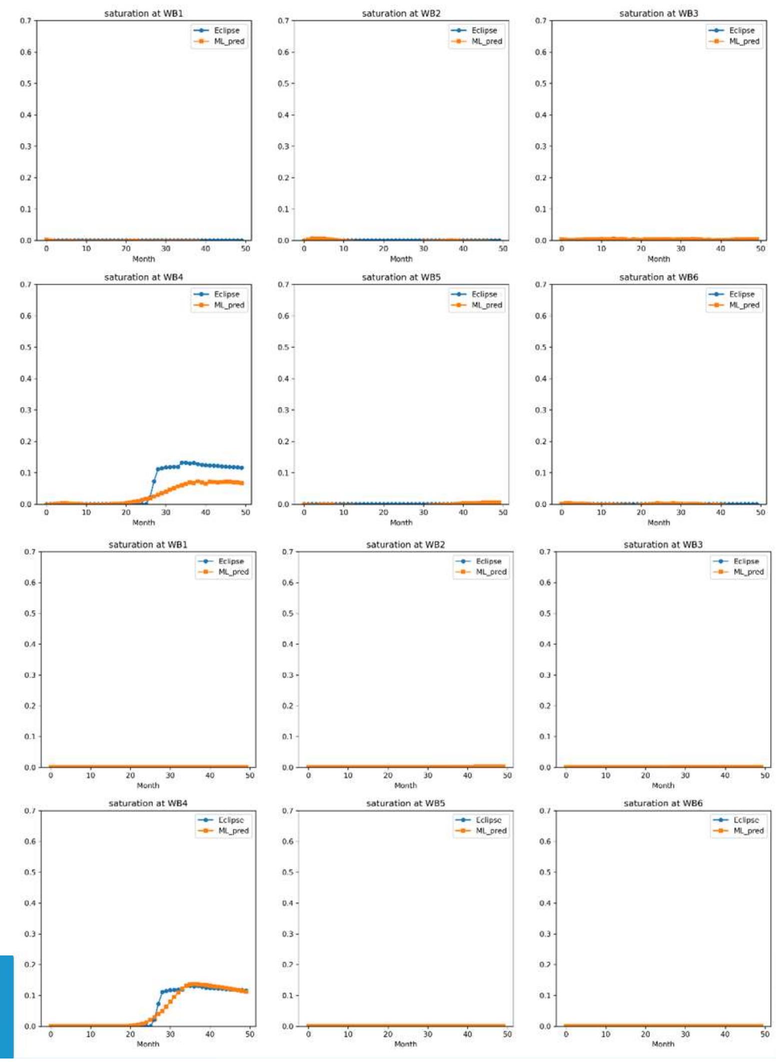
- Portions of this work were performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07-NA27344.
- Portions of this work were performed under the auspices of the U.S. Department of Energy by Pacific Northwest National Laboratory under contract DE-AC05-76RL01830.
- Portions of this work were performed under the auspices of the U.S. Department of Energy by Sandia National Laboratories under contract DE-NA0003525.
- Portions of this research were executed through the NETL Research and Innovation Center's Carbon Storage Program

Pressure (psi) & Saturation (-) at six different depths in monitoring well (realization 10)

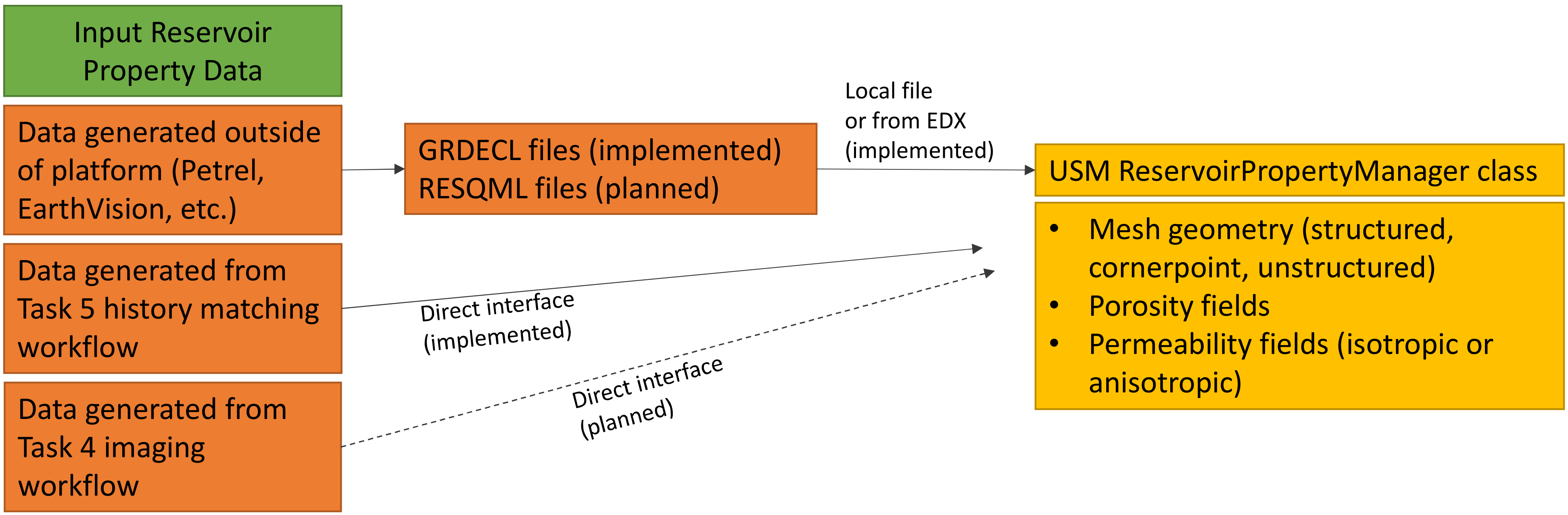


SNL

UT-BEG



Unified Simulation Module



Unified Simulation Module

Input Injection and Monitoring Well Data

Data generated outside of platform → *.csv, *.xlsx → USM OperationalScenarioManager class

Local file
or from EDX
(implemented)

Data generated from Task 5 history matching or VLE

Direct interface
(implemented
but not tested)

USM OperationalScenarioManager class

- Well location
- Perforation depths
- Injection rate
- Pressure

Unified Simulation Module

USM ReservoirPropertyManager

USM OperationalScenarioManager

USM ForwardModel classes

- Base class that defines the interface, making all forward models interchangeable
- Derived classes implement specific forward models
- Directly use ReservoirPropertyManager and OperationalScenarioManager objects
- Outputs ReservoirStateManagerObject

USM ReservoirStateManager

- Mesh geometry
- Pressure field time history
- Saturation field time history
- Stress field time history (planned)
- Strain field time history (planned)

Unified Simulation Module

USM ReservoirStateManager

- Mesh geometry
- Pressure field time history
- Saturation field time history
- Stress field time history (planned)
- Strain field time history (planned)

Output Dynamic Reservoir Data

Use outside of SMART platform enabled with:
- VTK files (implemented)
- RESQML files (planned)

Direct connection with other SMART platform components (VLE, history matching, etc.)