Development of a Framework for Data Integration, Assimilation, and Learning for Geological Carbon Sequestration (DIAL-GCS)

Project #: DE-FE0026515

Alex Sun, Ph.D., P.E.

Bureau of Economic Geology

The University of Texas at Austin

U.S. Department of Energy

National Energy Technology Laboratory

Addressing the Nation's Energy Needs Through Technology Innovation – 2019 Carbon Capture, Utilization, Storage, and Oil and Gas Technologies Integrated Review Meeting

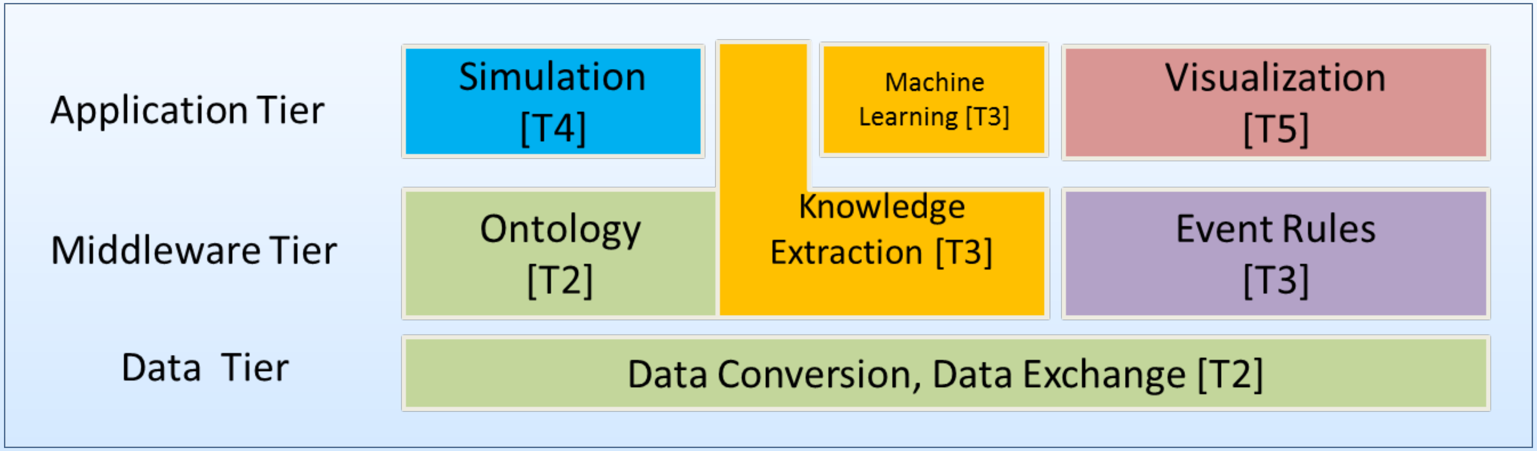
August 26-30, 2019

Presentation Outline

- Technical Status
  - Background and system design
  - Online anomaly detection using machine learning
  - Predictive modeling
- Accomplishments to Date
- Lessons Learned
- Synergy Opportunities
- Project Summary

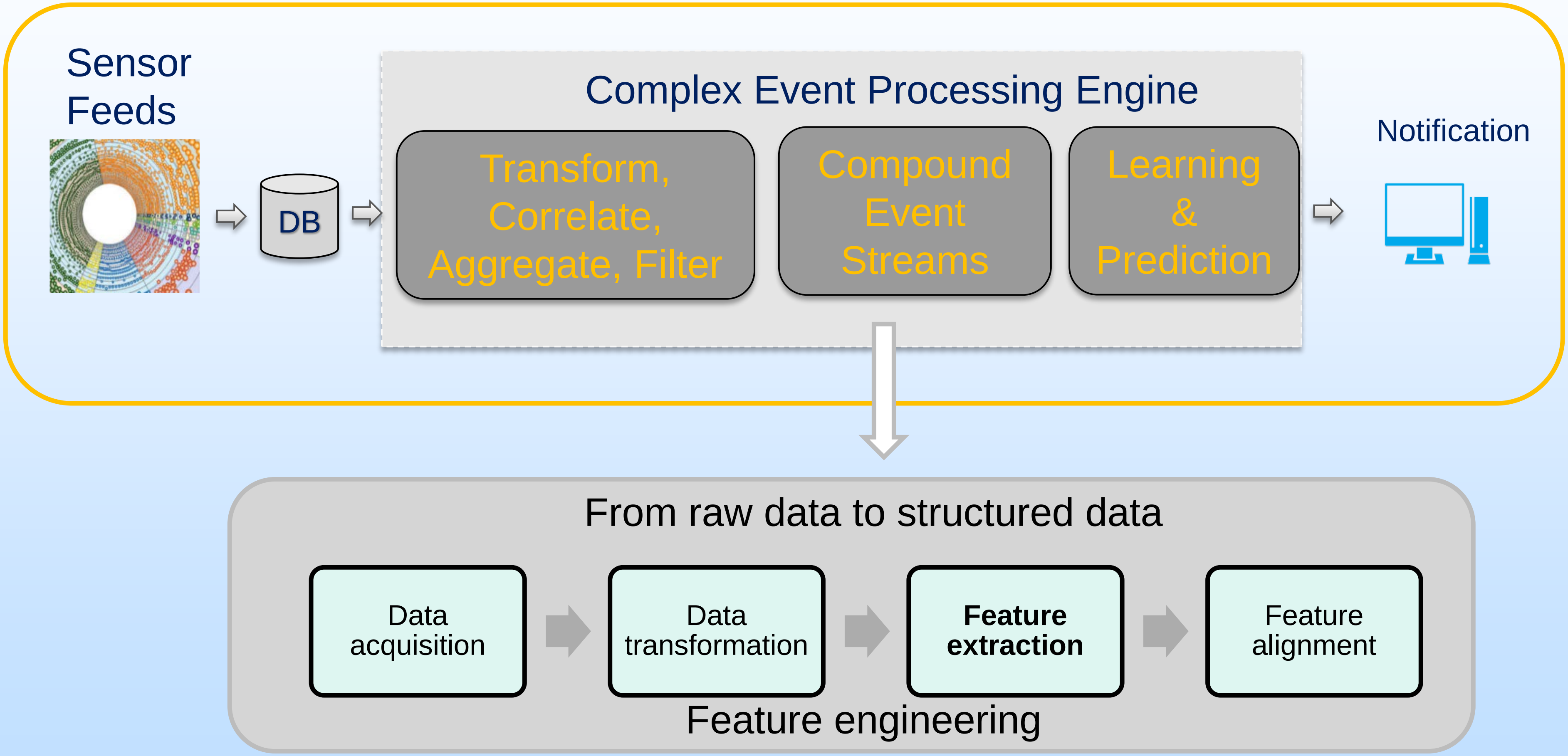
Background & Overview of Project

A multi-tier intelligent monitoring system (IMS)



Task 2: Sensor data schema development and provisioning(Y1)
Task 3: Development of CEP, machine learning (Y1-3)
Task 4: Coupled modeling, UQ, and data assimilation (Y1-4)
Task 5: System integration and demonstration (Y1-4)

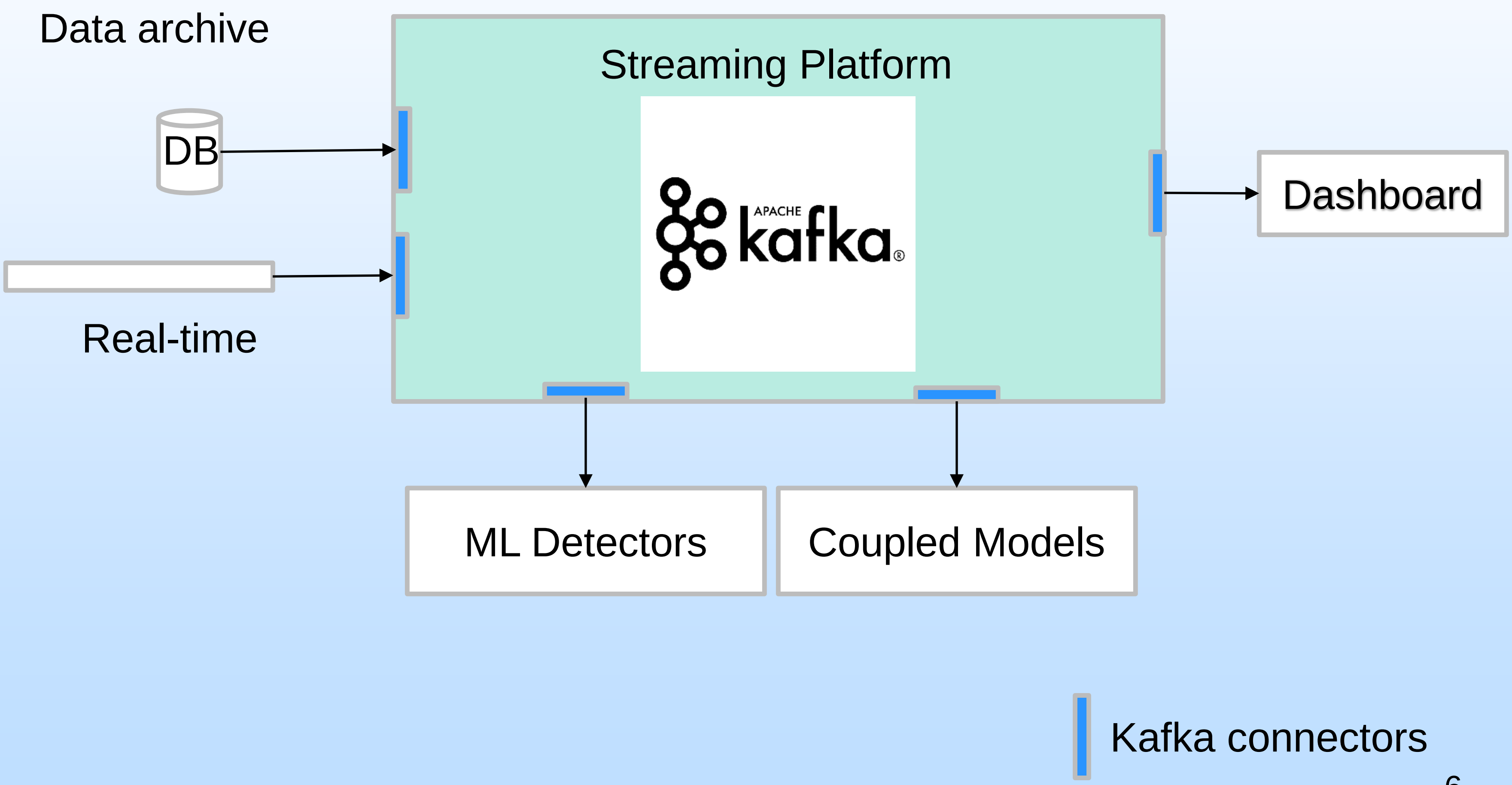
Task 3. Complex Event Processing



Data-Driven, ML-Enabled Anomaly Detection

- Machine learning (ML) is suitable for
  - Continuous monitoring
  - When physical process is not fully understood
  - Automated anomaly detection
- Requirements
  - Effective <u>online</u> ML algorithms
  - Labeled training data and expert insights!
  - High-performance, integrated computing infrastructure

Loosely Coupled Real-Time Data Processing System

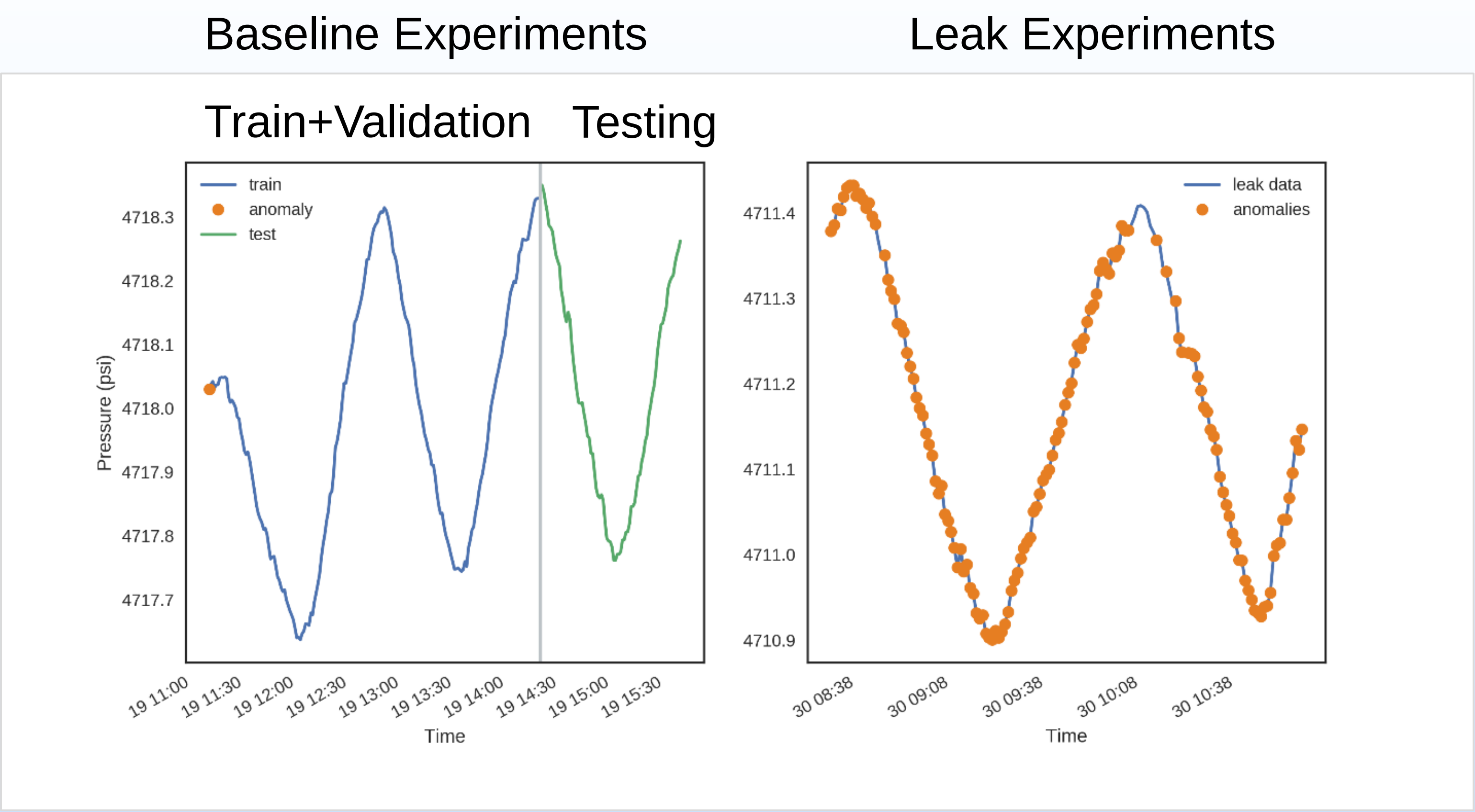


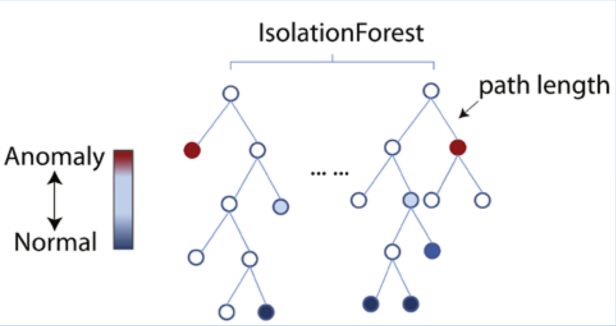
Anomaly Detection Case Study

Cranfield, MS experiments

Dataset include Pressure and Temperature measurements from
- Base pulse testing experiments (no known leaks)
- Controlled CO_2 release experiments (artificial leaks)

Anomaly Detection in Pressure Data



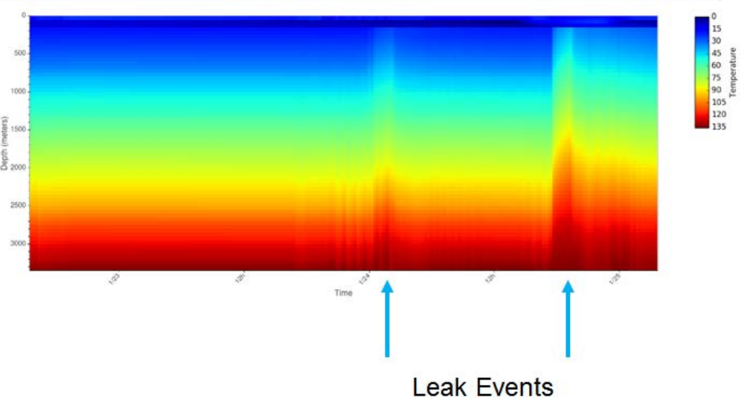


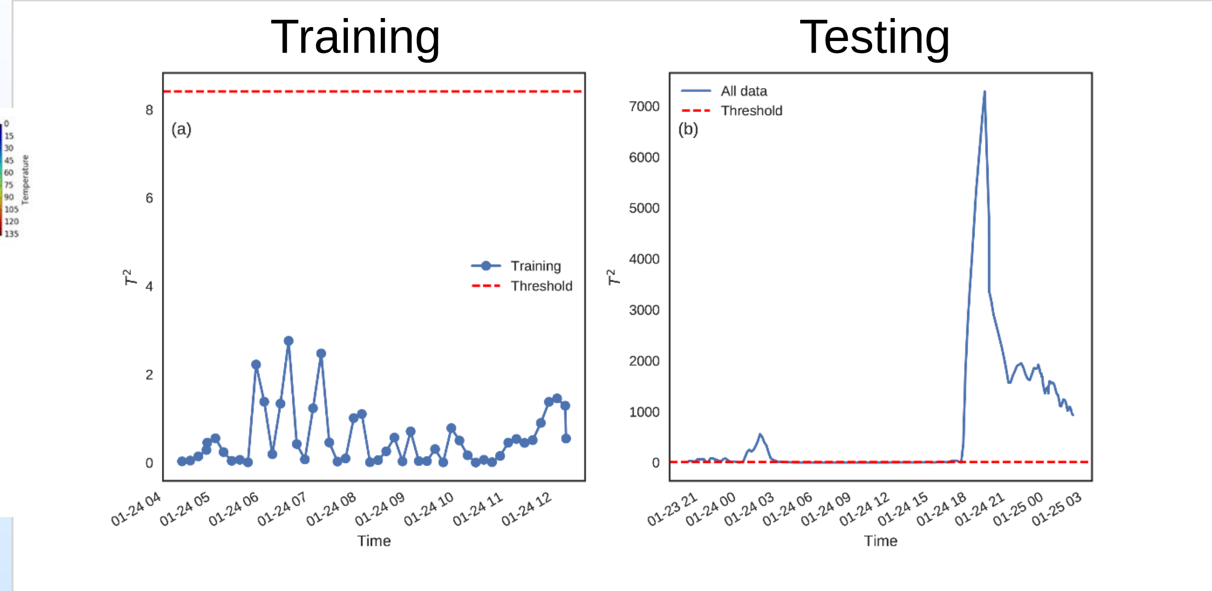
Use IsolationForest to perform unsupervised classification

- Randomly and recursively partition of data (trees)
- Calculate path length from root node to leaf node
- Use tree path length as a measure of normality
- Outliers correspond to shorter trees

Anomaly Detection in DTS Data

DTS Data





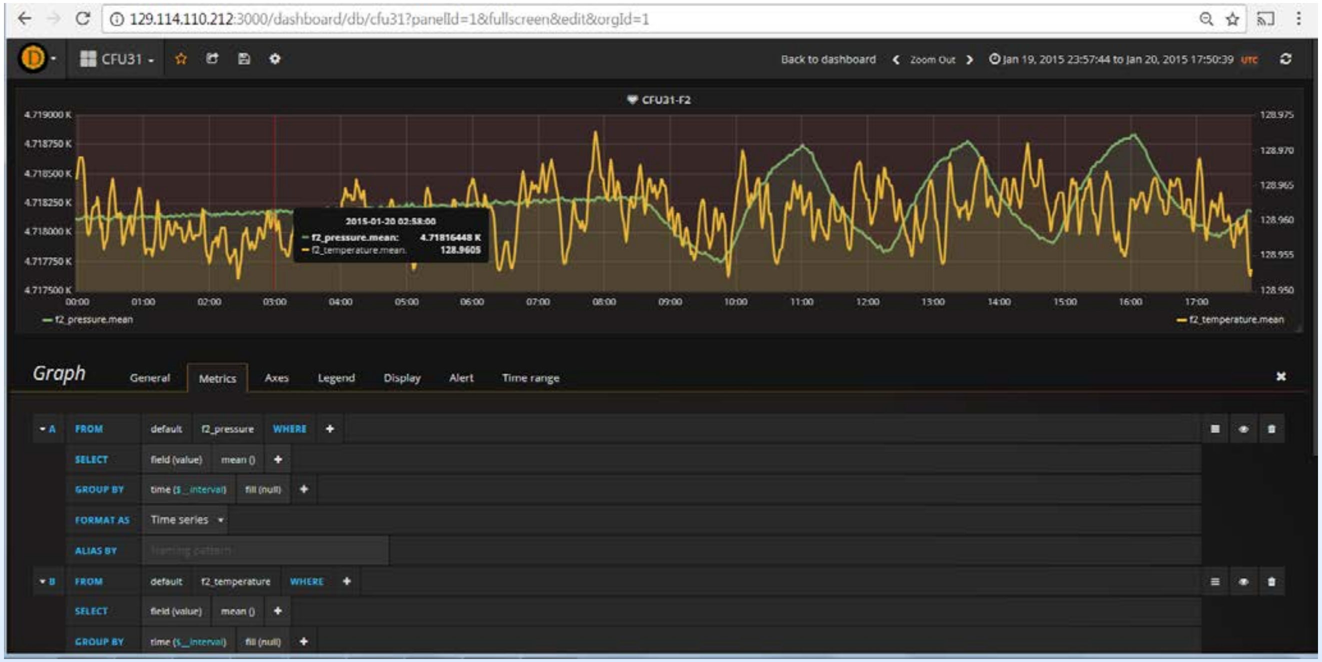
Calculate T^2 statistic

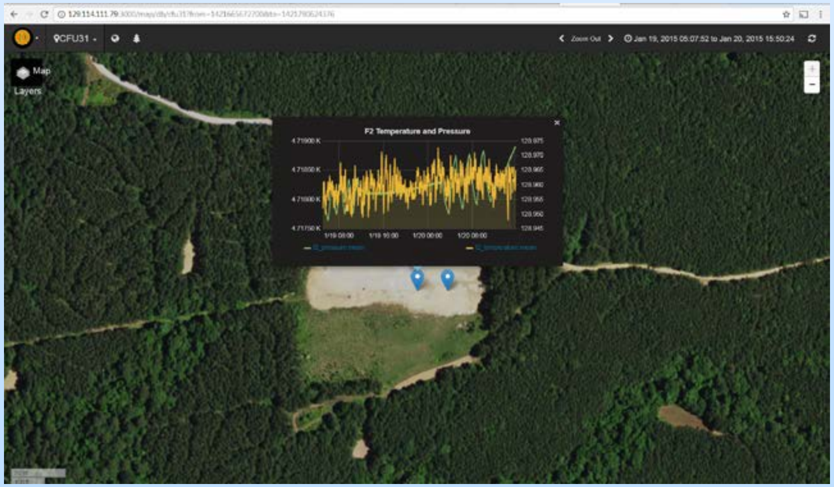
- Form normalized training sample matrix
- Perform PCA (data compression) on the sample matrix
- Calculate *T^2 statistic* for a user-defined significance level

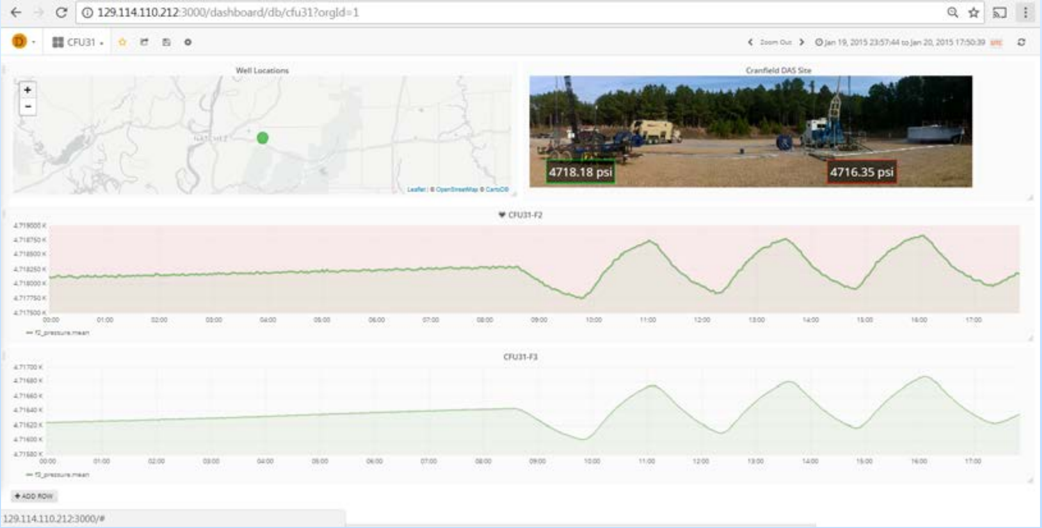
Sun et al., Env Modeling & Software, 2019
Zhong et al., 2019b

DIAL-GCS 1.0

Design 1.0:
- Web GIS
- Time series management
- A lot custom coding



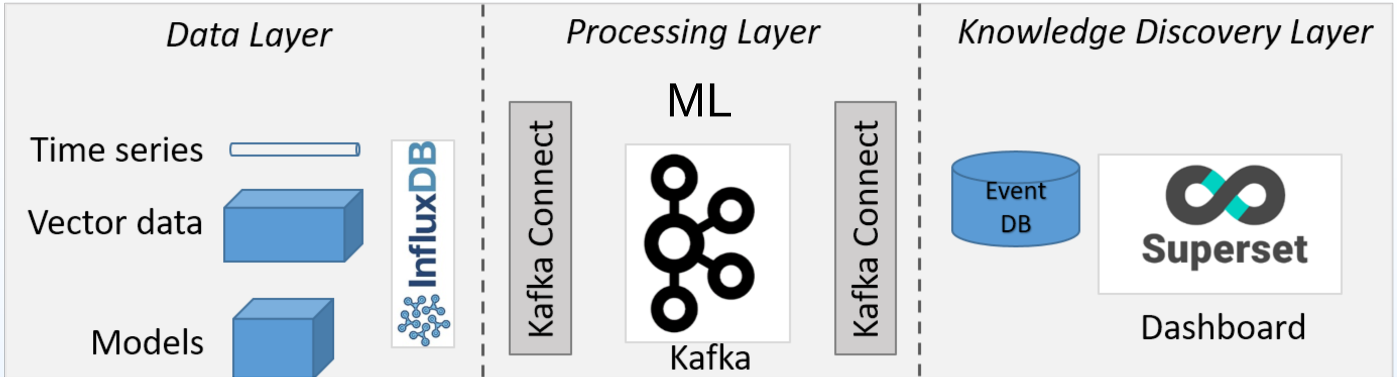


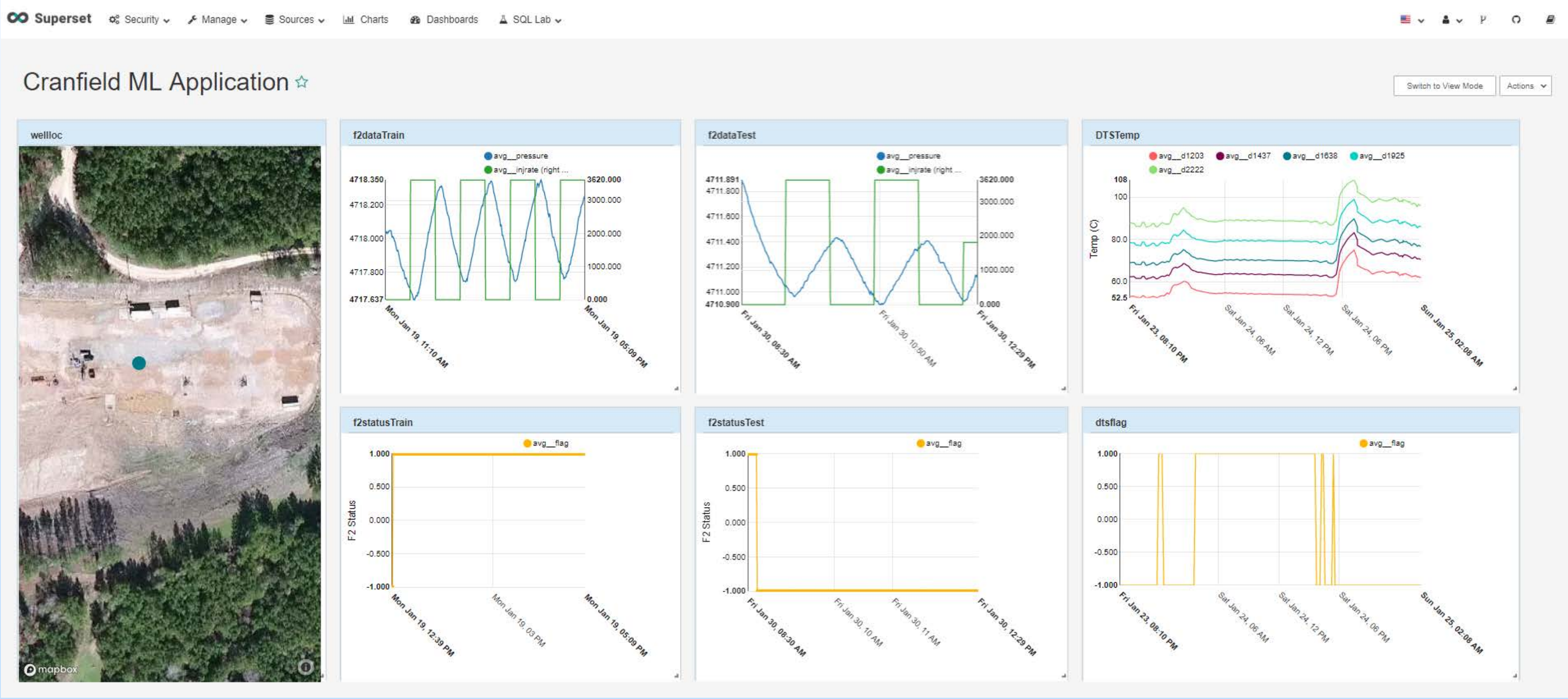


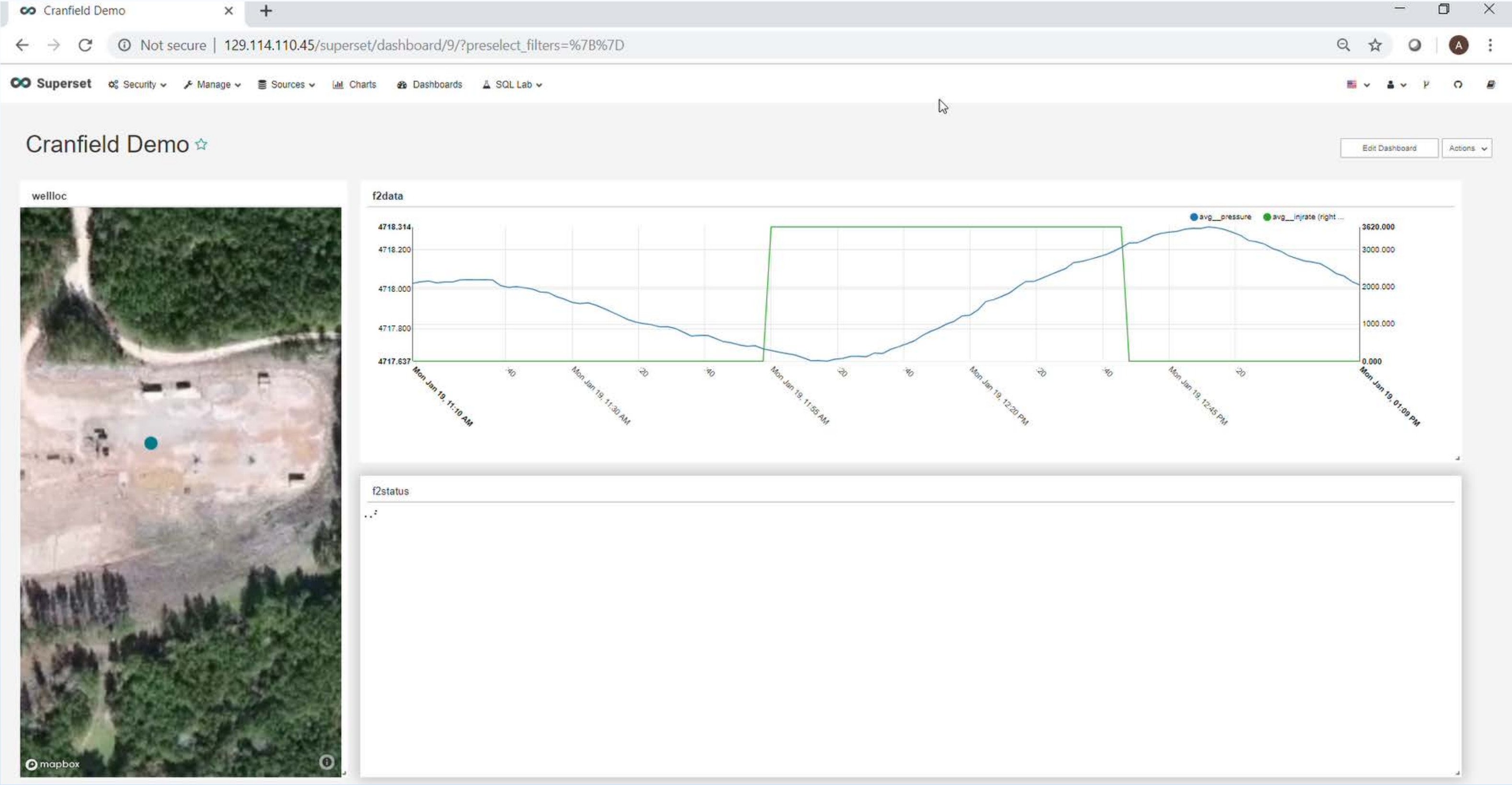
DIAL-GCS 2.0

Design 2.0:

- Kafka-based
- Flexible



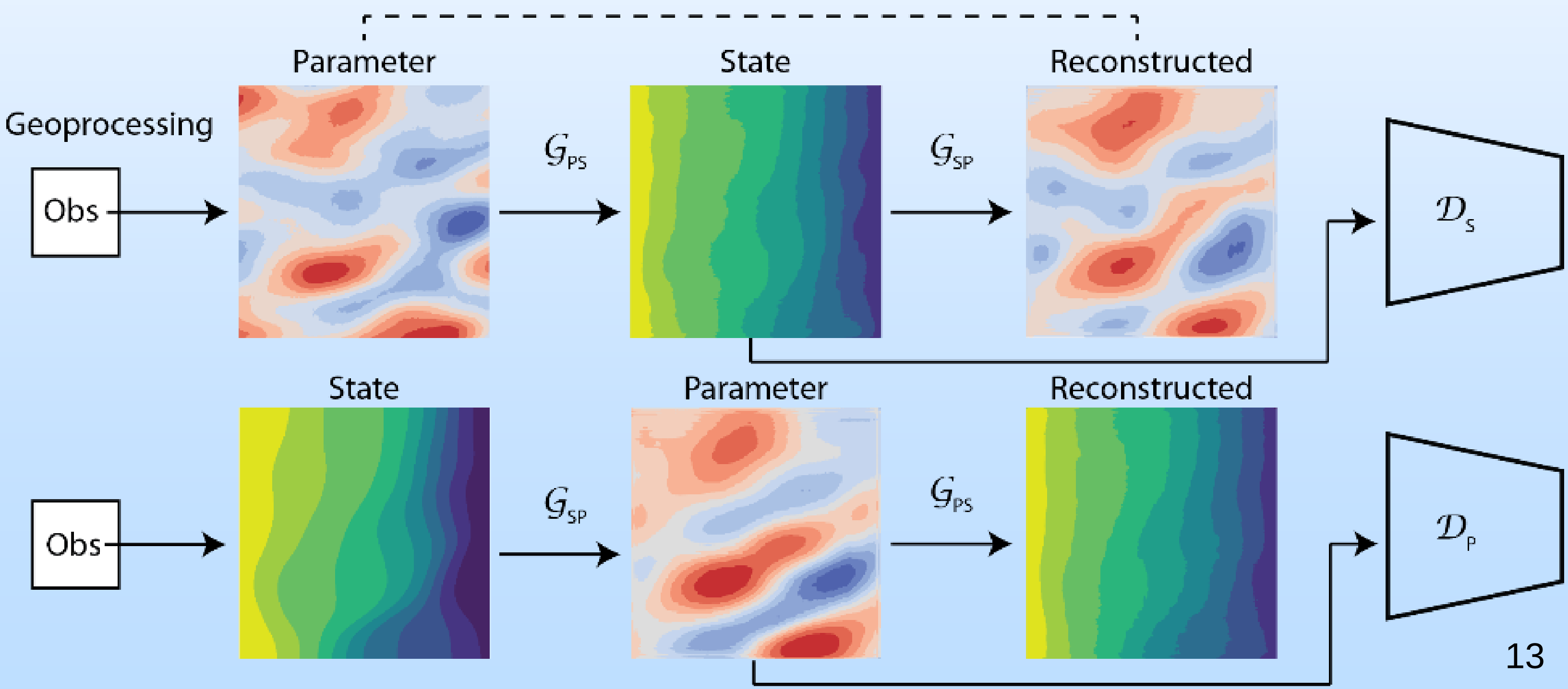



Cranfield Demo
Not secure | 129.114.110.45/superset/dashboard/9/?preselect_filters=%7B%7D
Superset
Security
Manage
Sources
Charts
Dashboards
SQL Lab
Cranfield Demo
Edit Dashboard
Actions
wellloc
f2data
f2status
mapbox
4718.314
4718.200
4718.000
4717.800
4717.637
3620.000
3000.000
2000.000
1000.000
0.000
Mon Jan 19, 11:10 AM
Mon Jan 19, 11:30 AM
Mon Jan 19, 11:55 AM
Mon Jan 19, 12:20 PM
Mon Jan 19, 12:45 PM
Mon Jan 19, 01:09 PM

Task 4. Deep Learning Based Surrogate Modeling

- Deep learning (DL) is a very powerful tool for pattern recognition.
- In geosciences, there's a lot of hype on DL but also many questions
- We developed an innovative DL pipeline for combining DL with physics-based models

A generic simulation/inversion framework

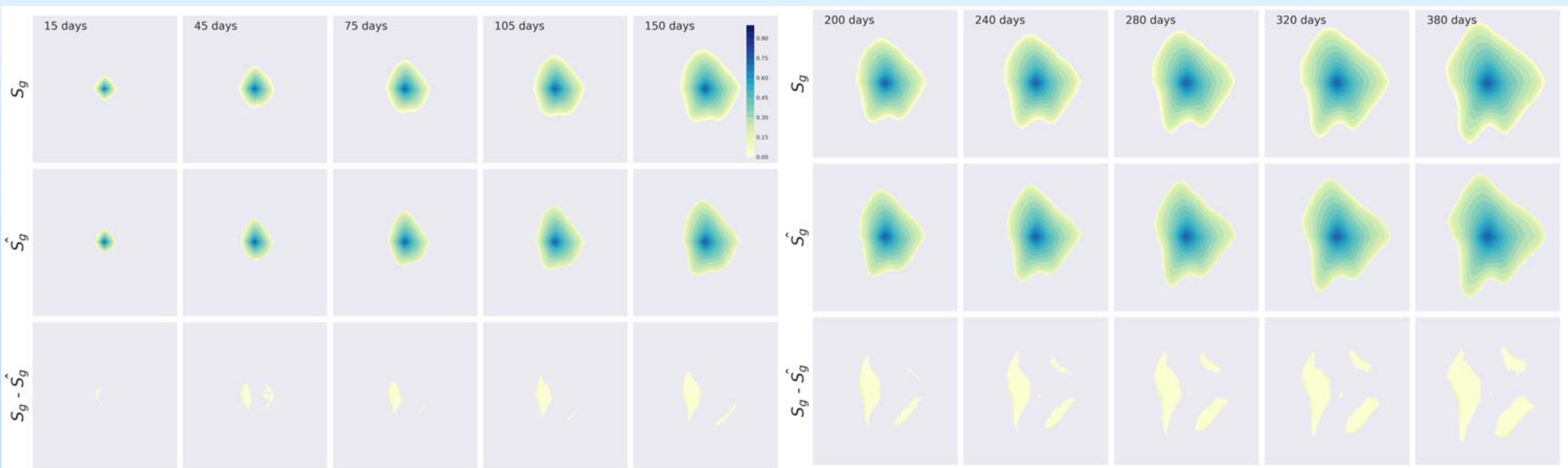


Sun 2018 GRL

Deep learning based surrogate models for CO_2 plume prediction

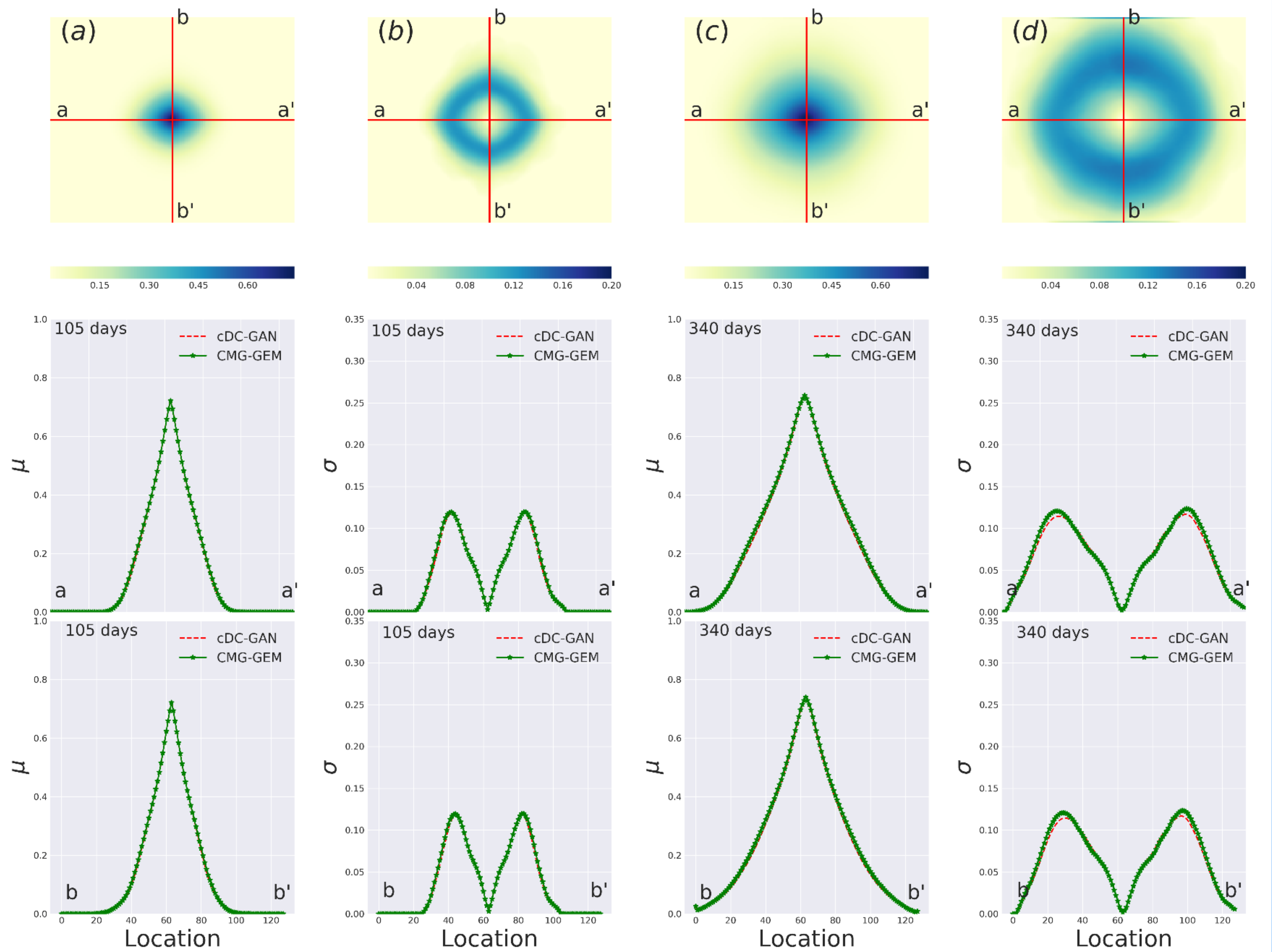
Input: Permeability field
Output:
Sg: Actual CMG-GEM result
$\hat{S}g$: Surrogate model results

Once trained, our model can be used to predict CO2 plume for given input properties and at any time

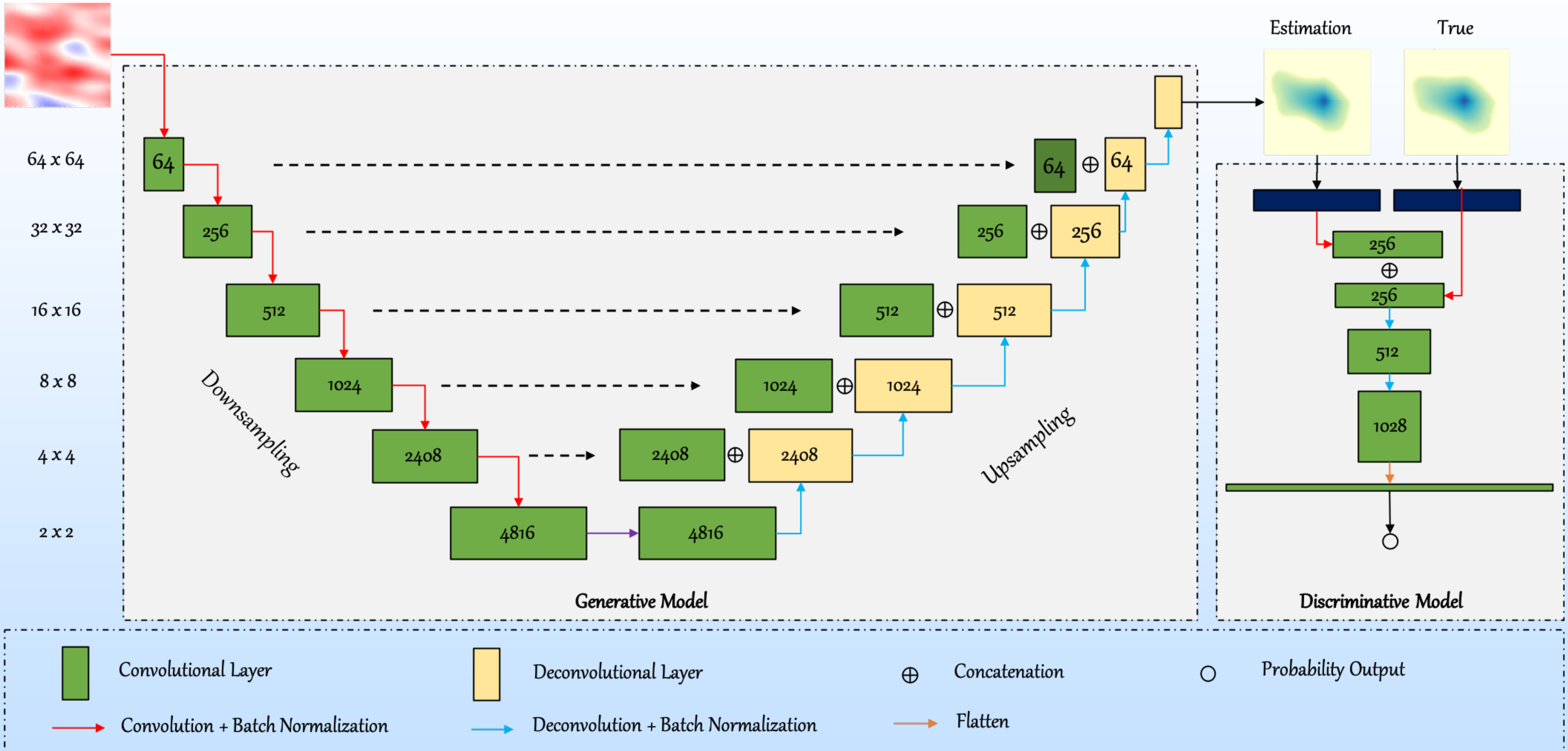


Time

Validation: Comparison with Monte Carlo Simulation



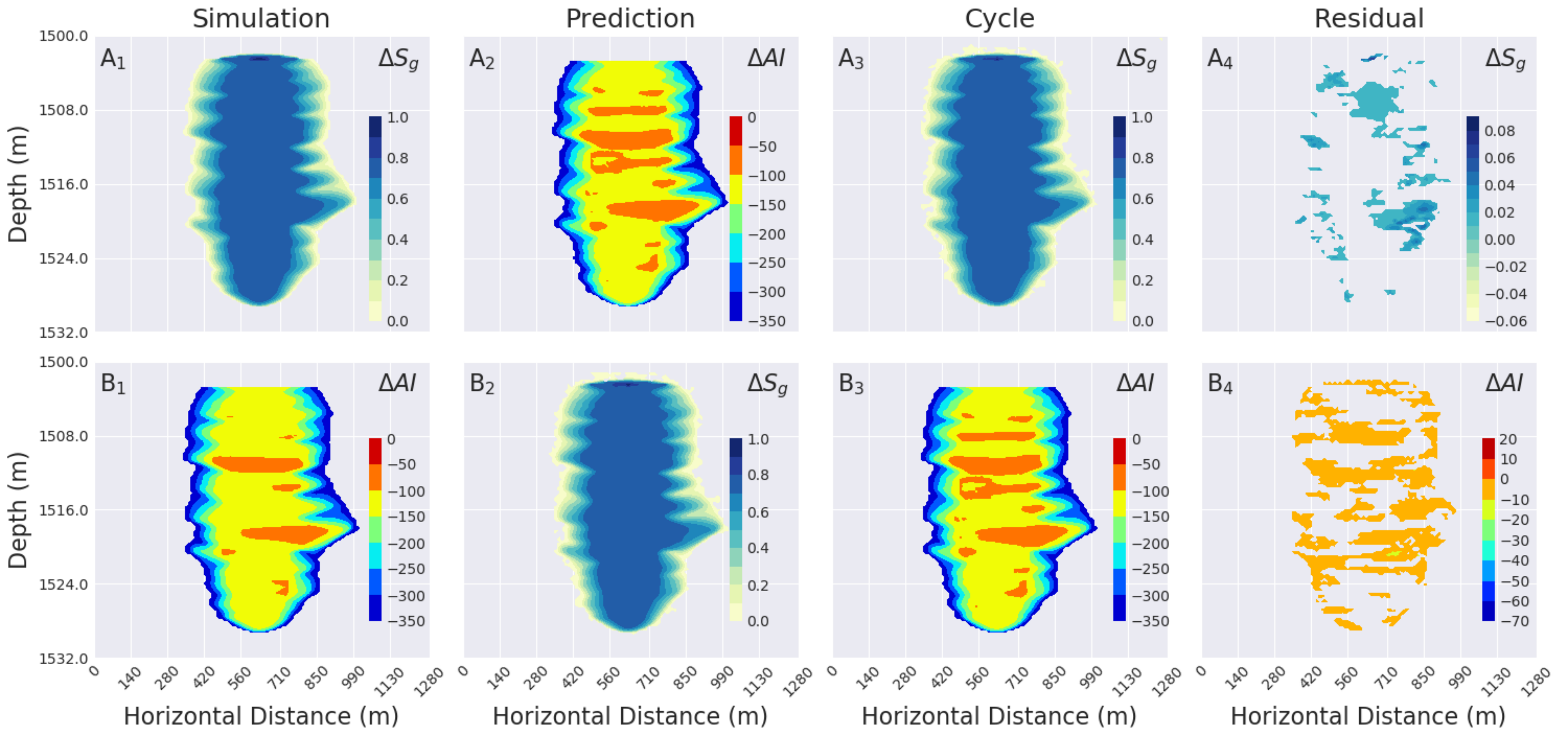
Dynamic mapping: Prediction of CO_2 Plume Movement



Zhong et al., 2019a

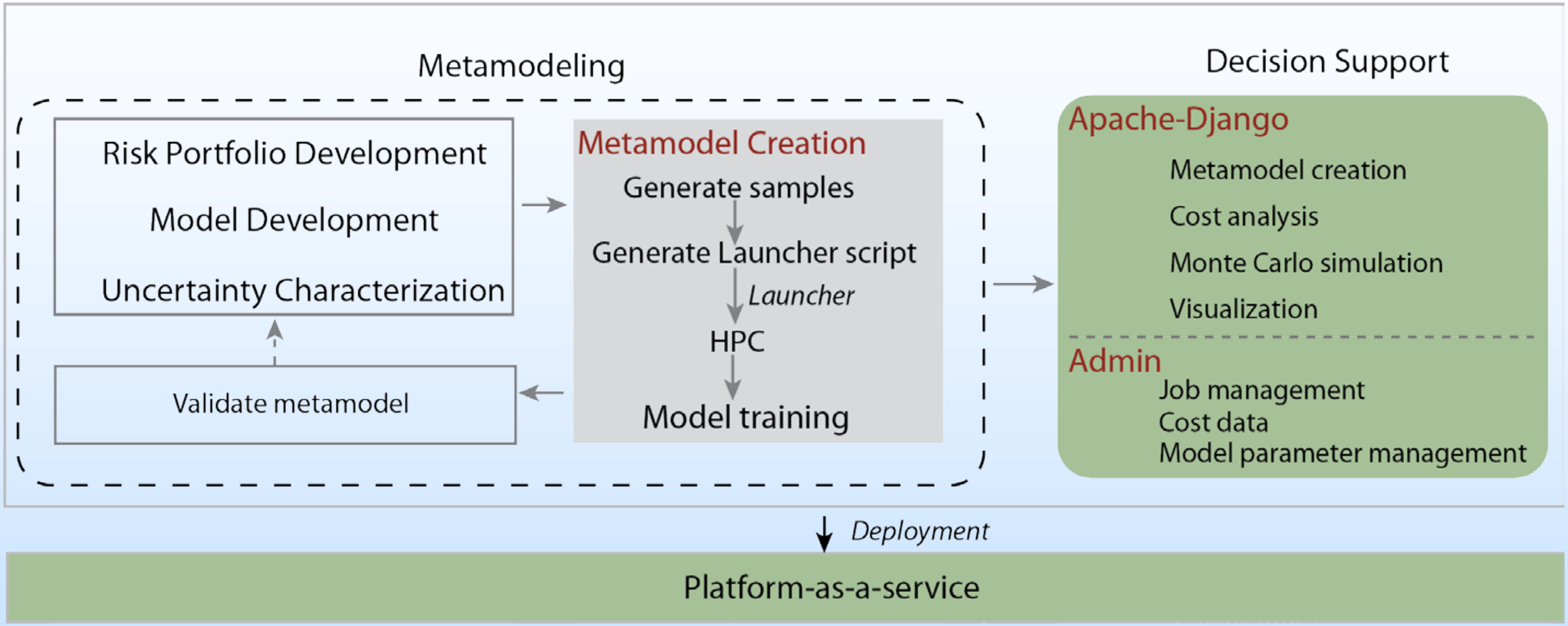
Combining with Geophysics Data

Saturation ⟷ Acoustic Impedance



Zhong et al., under review

Web-Based Monitoring Planning



Sun et al., 2018, Metamodeling-based approach for risk assessment and cost estimation: Application to geological carbon sequestration. Computers & Geosciences.

Leakage Assessment and Cost Estimation Tool

Admin | Metamodeling | Cost Estimate |

Browse... No file selected. Upload

File uploaded at: media/metamodel.json

Variable	Distribution	Parameters
Log Reservoir Permeability:	normal	[-29.93360621, 0.5]
Log AZMI Permeability:	normal	[-30.62675339, 0.5]
Reservoir Porosity:	uniform	[0.1, 0.2]
AZMI Porosity:	uniform	[0.05, 0.3]
Aquitard thickness:	uniform	[10.0, 30.0]
Injection Rate:	uniform	[0.5, 5]

Algorithm: gpr

Create Metamodel

Reservoir porosity[-]:	0.2
Aquifer porosity[-]:	0.1

Types of metamodeling supported:

- Gaussian process regression
- Sparse grid

Permeability (m^2):	0.00000000000005
Porosity (-):	0.2
Injection Rate [Mt/yr]:	3

Above Zone Parameters

CO2 density (kg/m^3):	479
Brine density (kg/m^3):	1045
CO2 viscosity (Pa*s):	0.0000395

y [m]

-1,

-2,

-3,

Total le

Estima

istration

WELCOME, ADMIN. VIEW SITE / CHANGE PASSWORD / LOG OUT

Home › Reducedmodel › Expert elicitations

Select expert elicitation to change

ADD EXPERT ELICITATION +

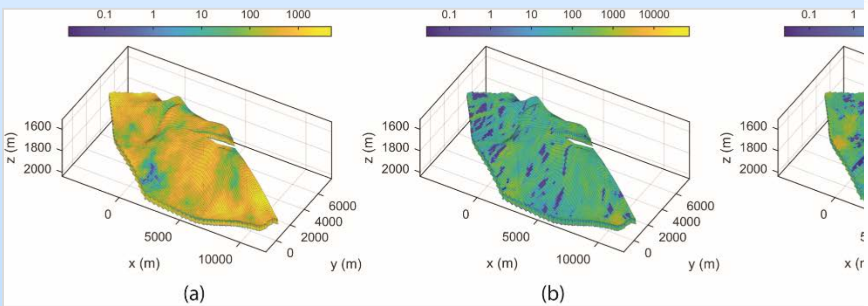
Action: --------- Go 0 of 32 selected

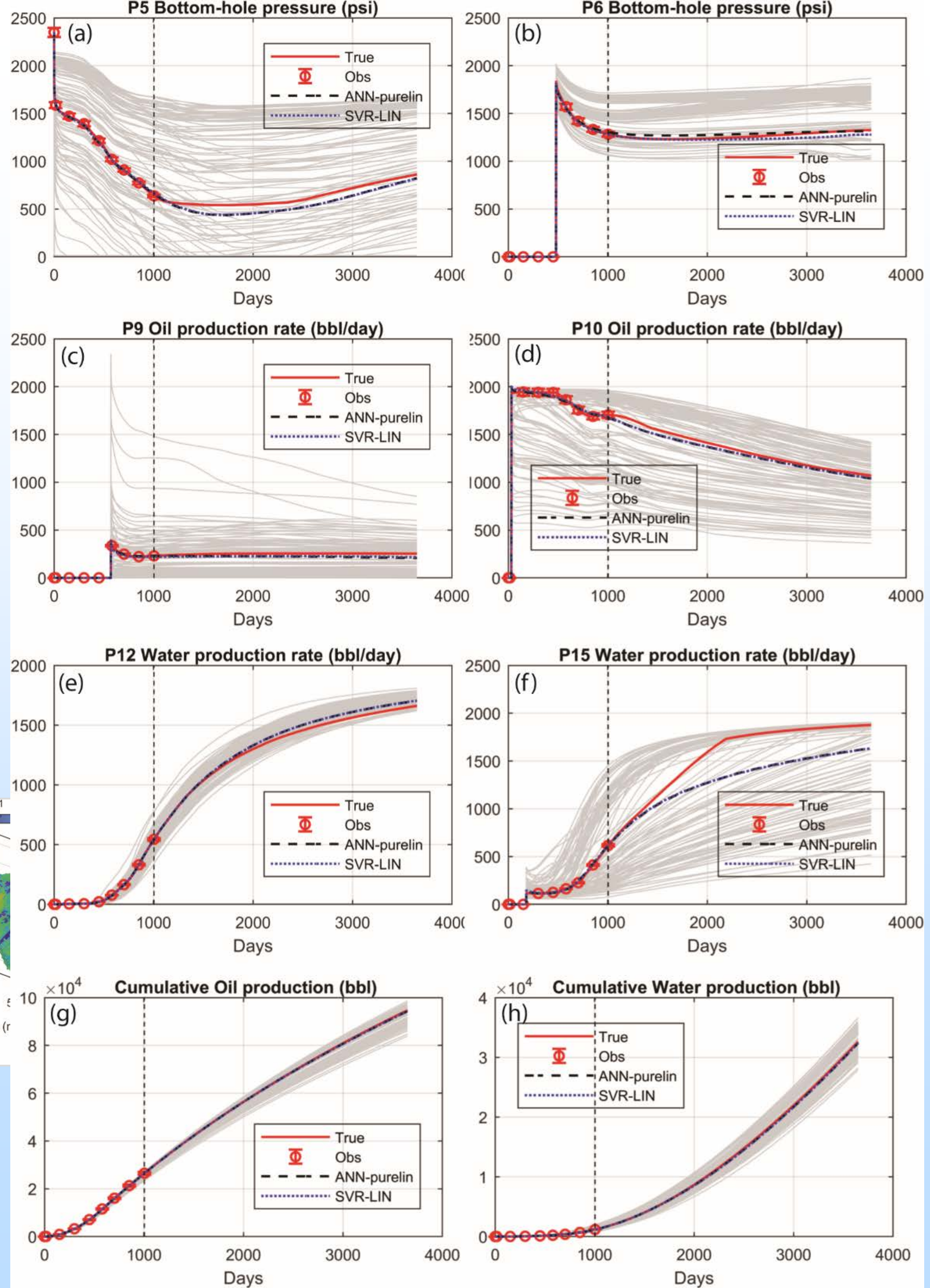
EXPERT	LEAK COST CATEGORY	SCALE
Teresa C.	B: Injection interruption	1
Teresa C.	B: Legal costs	3
Teresa C.	B: Well remediation	6
Teresa C.	C: Other property interference	1
Teresa C.	C: Injection interruption	3
Teresa C.	C: Environmental remediation	2
Teresa C.	C: Legal costs	1
Teresa C.	C: Well remediation	3

Data-Space Inversion (DSI)

What is DSI?

- A new paradigm for long-term prediction and UQ without using history matching
- Prior knowledge is used to generate possible scenarios, but not to calibrate model
- DSI combines physically-based model with ML





Jeong et al., 2018a, A learning-based data-driven forecast approach for predicting future reservoir performance. AWR.

Optimal Monitoring Network Design

Objective Function

Well cost =
CAPEX($/well)+
OPEX($/well/day)+
Intervention($/well)

Leakage cost =
Brine($/ton)+
CO_2($/ton)

Optimization toolbox

Binary Integer Programming

- Linear problem
- Convex

Optimize monitoring network

Constraints

of monitoring wells ≤ N_{max}

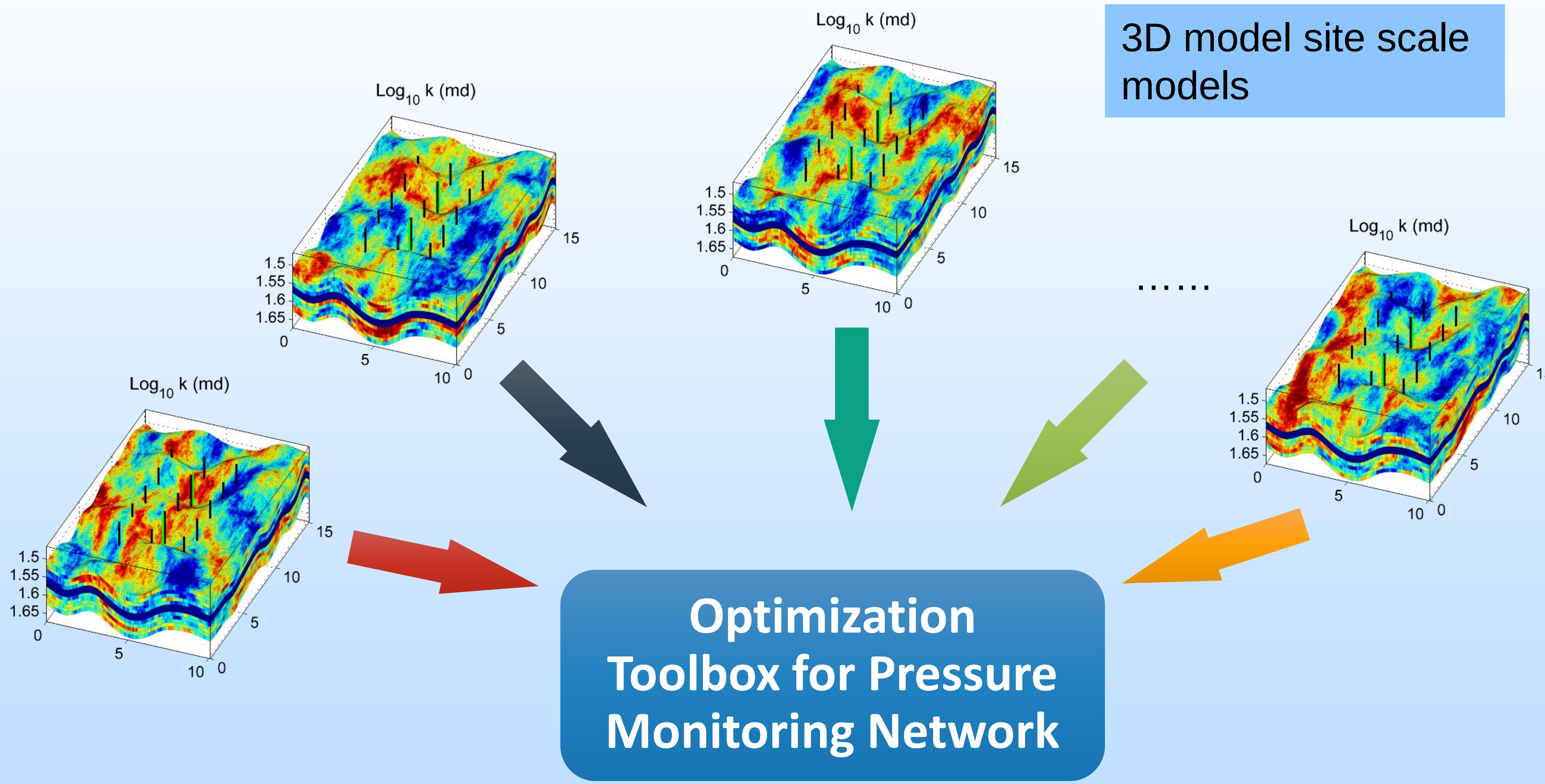
CO_2 leakage ≤ M% of total injected CO_2

ΔP at $t_{\text{leakage detection}} \geq \Delta P_{\text{threshold}}$

Jeong et al., 2018b, Cost-optimal design of pressure-based monitoring networks for carbon sequestration projects, with consideration of geological uncertainty, International Journal of Greenhouse Gas Control.

Our tool maximizes NPV by considering

- High uncertainty in geologic models
- Monitoring budget
- Leakage damage cost
- Carbon credit *: 45Q Tax Incentives for CCUS*



Lessons Learned

- We are developing an intelligent monitoring system to help extract intelligent information. Our applications include
  - Web-based monitoring planning
  - Pressure-based monitoring network design
  - Data space inversion
  - Deep learning tools
- Data-driven machine-learning models are suitable for continuous monitoring and anomaly detection and can be used together with physics-based models for surrogate modeling
- A viable approach is to combine prior information, expert knowledge, and state-of-the-art machine learning tools for knowledge discovery and representation

Accomplishments to Date

- Task 2: Data management
  - Year 1: Developed schema and data adaptors for storing, exchanging information, and visualizing information
- Task 3: Complex event processing using machine learning
  - Year 2: Implemented predictive models on different test datasets
  - Year 3-4: Updated the existing platform for usability
- Task 4: Coupled modeling / data assimilation
  - Year 2: Implemented workflow for automating data assimilation. Demonstrated Web-based modeling approaches
  - Year 3-4: Focused on ML and DL tool development
- Task 5: Integration and demonstration
  - Year 1-4: Experimented with a large number of web-based technologies for making the system more user friendly

Synergy Opportunities

- DIAL-GCS is an intelligent monitoring system designed for anomaly detection, monitoring network design, physics-based machine learning, leakage cost estimation
- Most tools are web-based, or can be readily converted to web-based, for CCS decision support needs

Project Summary

- Developed and improved DIAL system
- All tasks are on revised schedule
- Next steps
  - Formalize data transformation and work flow
  - Provide deep learning based web service
  - Integrate different analytic modules and disseminate results
  - Wrap the project in the next year

Acknowledgements

- DOE/NETL PM: Bruce Brown

- University of Texas
  - Bureau of Economic Geology: Sue Hovorka, Katherine Romanak, Hoonyoung Jeong, Zhong Zhi
  - Texas Advanced Computing Center

- LBNL: Barry Freifeld (provided DTS data)

Appendix

– These slides will not be discussed during the presentation, but are mandatory.

Benefit to the Program

- Carbon storage program goals being addressed

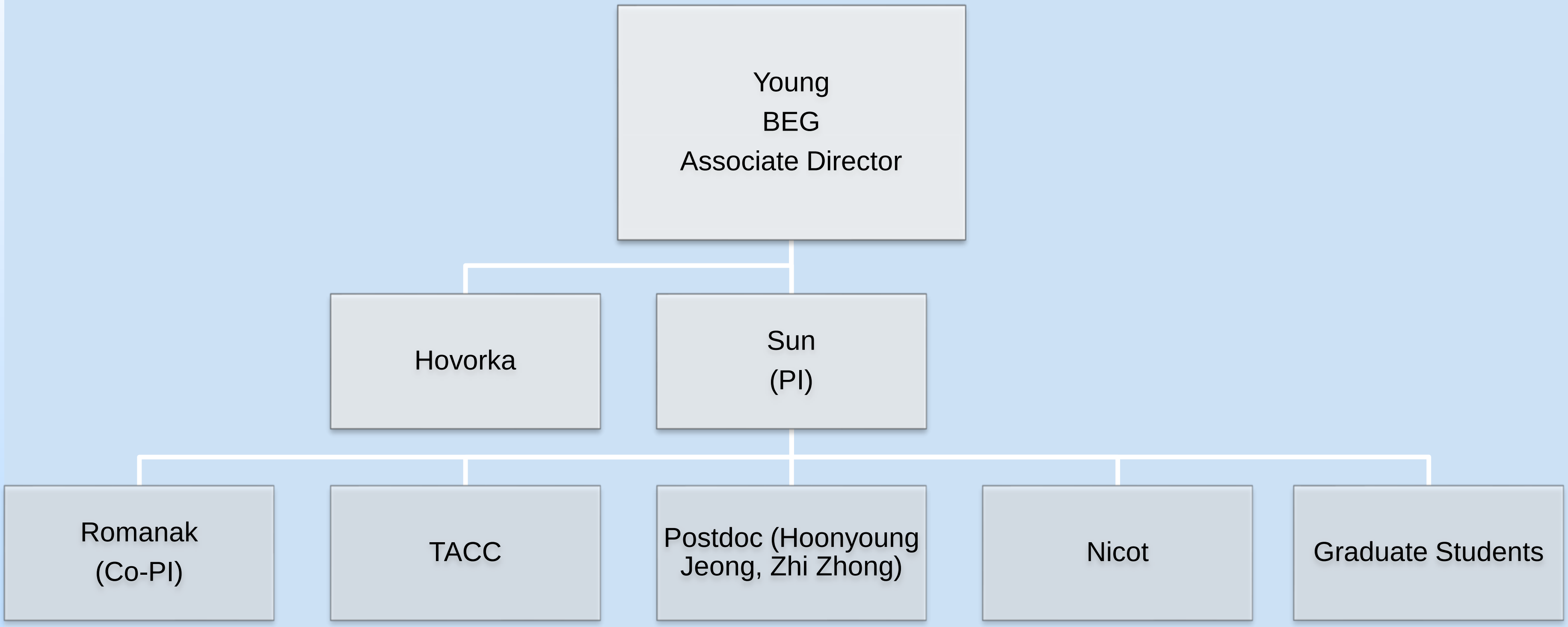
Develop and validate technologies to ensure 99 percent storage permanence

- Expected benefits of this IMS Project
  - Transform scientific knowledge to decision power and public knowledge
  - Promote data sharing and visual analytics
  - Better collaboration among team members
  - Public outreach
  - Streamline CCS data management and decisionmaking
  - Facilitate the optimal allocation of monitoring resources

Project Overview
Goals and Objectives

- Develop GCS data management module for storing, querying, exchanging, and visualizing GCS data from multiple sources and in heterogeneous formats
  - Success Criterion: Whether a flexible, user-friendly Web portal is set up for enabling data exchange and visual analytics
- Incorporate a complex event processing (CEP) engine for detecting abnormal situations by seamlessly combining expert knowledge, rule-based reasoning, and machine learning
  - Success Criterion: Whether a set of decision rules are developed for identifying abnormal signals in monitoring data
- Enable uncertainty quantification and predictive analytics using a combination of coupled-process modeling, data assimilation, and reduced-order modeling
  - Success Criterion: Whether a suite of computational tools are developed for UQ and predictive analytics
- Integrate and demonstrate the system's capabilities with both real and simulated data
  - Success Criterion: Whether the IMS tools developed under Goals A to C are integrated, streamlined, and demonstrated for a realistic GCS site

Organization Chart



Gantt Chart

Table 1. Revised Project Gantt chart
(Numbers in table rows indicate milestones).
(BP1-2 ; NCTE)

Task	Description	Year 1				Year 2				Year 3				Year 4				Year 5			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Update project management plan																				
2	Sensor data management																				
2.1	Ontology/schema development																				
2.2	Sensor data adaptor development																				
3	CEP Development																				
3.1	Rule definition																				
3.2	Reasoning and learning																				
3.3	Testing																				
4	Coupled modeling/Assimilation																				
4.1	Coupled modeling																				
4.2	Reduced order modeling																				
5	Integration and demonstration																				
5.1	Integration																				
5.2	Demonstration																				
6	Synthesis of results																				
6.1	Dissemination of results																				
6.2	Technology transfer																				

Bibliography

– Peer-Review Manuscripts

- Sun, A., Z. Zhong, H. Jeong, and Q. Yang, 2019, Building complex event processing capability for intelligent environmental monitoring, Environmental Modeling & Software, 116, 1-6.
- Zhong, Z., Sun, A. Y., & Jeong, H. 2019a, Predicting CO2 plume migration in heterogeneous formations using conditional deep convolutional generative adversarial network. Water Resources Research.
- Zhong, Z., Sun, A. Y., Yang, Q., & Ouyang, Q., 2019b, A deep learning approach to anomaly detection in geological carbon sequestration sites using pressure measurements. Journal of Hydrology, 573, 885-894.
- Sun, A. Y., 2018, Discovering state-parameter mappings in subsurface models using generative adversarial networks, Geophysical Research Letters, 45(20), 11,137-11,146.
- Sun, A. Y., Jeong, H., Gonzalez, A., and Templeton, T., 2018, Metamodeling-based approach for risk assessment and cost estimation: application to geological carbon sequestration, Computers and Geosciences, v. 113, p. 70-80.
- Jeong, H., Sun, A. Y., Lee, J., and Min, B., 2018a, A learning-based data-driven forecast approach for predicting future reservoir performance. Advances in Water Resources, v. 118, p. 95-109.
- Jeong, H., Sun, A. Y., and Zhang, X., 2018b, Cost-optimal design of pressure-based monitoring networks for carbon sequestration projects, with consideration of geological uncertainty, International Journal of Greenhouse Gas Control, v. 71, p. 278-292.

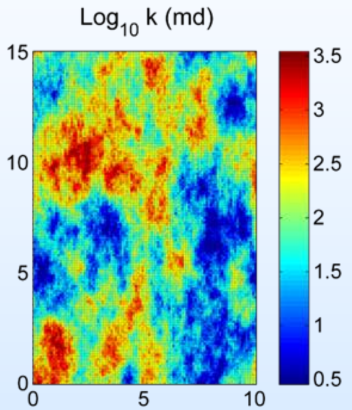
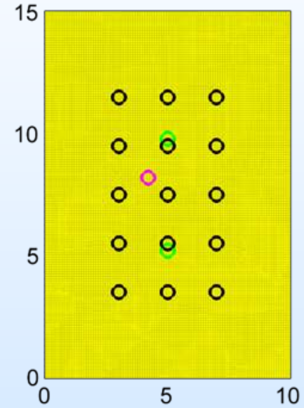
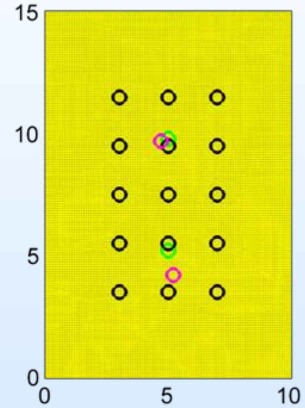
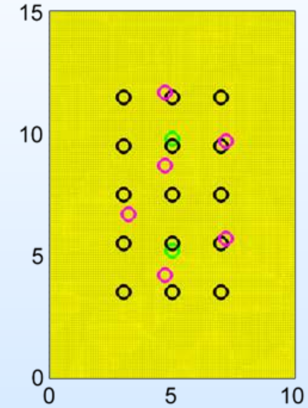
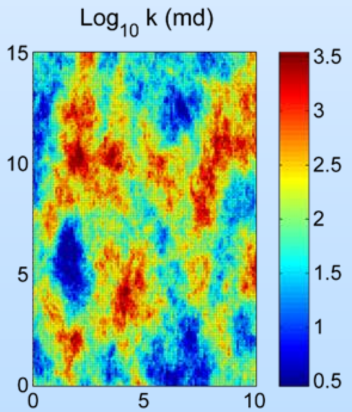
– Presentations

- Development of anomaly detection models for deep subsurface monitoring, presented at the fall meeting of American Geophysical Union, New Orleans, LA, December, 2017

The optimal monitoring well locations are different because heterogeneous permeability affects

- Spatial pressure distribution
- Leakage detection time

Black: leaky well
Green: injector
Magenta: monitoring well

		C_{brine}	C_{CO2}	C_{brine}	C_{CO2}	C_{brine}	C_{CO2}
Geologic model		**$10/t**	**$10/t**	**$10/t**	**$1,000/t**	**$100/t**	**$10/t**
	**Total cost**	$8.76 MM		$9.63 MM		$29.75 MM	
	Optimal monitoring well location						
	**Total cost**	$9.16 MM		$9.99 MM		$31.37 MM	
	Optimal monitoring well location